DEVELOPMENT AND SIGNALING: STUDIES OF SOXC, RGM AND NEOGENIN HOMOLOGS IN C. ELEGANS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Chenxi Tian

January 2013

DEVELOPMENT AND SIGNALING: STUDIES OF SOXC, RGM AND NEOGENIN HOMOLOGS IN C. ELEGANS

Chenxi Tian, Ph. D.

Cornell University 2013

A fundamental question in developmental biology is how a single multipotent cell undergoes cell divisions and fate specification to give rise to different cell types. In this process, multiple signaling events and transcription factors play critical roles. The understanding of the developmental mechanisms is especially important because malfunction of many developmental players usually associates with various human diseases.

I have studied a transcription factor SEM-2, a *C. elegans* homolog of vertebrate SoxC proteins. Mis-regulation of SoxC has been shown to associate with multiple types of tumors. However, the molecular mechanism by which SoxC gene expression is regulated is not known. I have determined the function of the sole *C. elegans* SoxC homolog SEM-2 in the M lineage, which produces the postembryonic mesoderm. I found that SEM-2/SoxC is both necessary and sufficient to promote a proliferating blast cell fate, the sex myoblast fate, over a differentiated striated bodywall muscle fate. I have shown that this function of *sem-2* is directly regulated by PBC/HOX factors in the M lineage, as demonstrated by computational, *in vivo* functional and *in vitro* electrophoretic mobility shift assays (EMSA). I also identified the positional cues that dictate the specific expression of *sem-2* in the sex myoblast precursors and their descendants, which include BMP signaling, Notch signaling and Wnt signaling.

Ongoing studies include identifying and analyzing the direct downstream targets of *sem-2* in promoting the proliferative over differentiative fates by a candidate gene approach.

Another focus of my thesis is on two modulators of BMP signaling in *C. elegans*: DRAG-1, a GPI (glycophosphatidlinositol)-anchored protein that is the sole member of the repulsive guidance molecule (RGM) family of proteins, and UNC-40, a transmembrane protein that is the sole *C. elegans* DCC(Deleted in Colorectal Cancer)/neogenin ortholog, best studied for its roles in axon and cell movement. The vertebrate homologs of both proteins have been found to modulate BMP signaling, but the mechanistic details of this regulation are not well understood. Using a combination of molecular genetic and biochemical analyses, I demonstrated that both DRAG-1/RGM and UNC-40/neogenin are positive modulators of BMP signaling at the ligand-receptor level. The role of UNC-40/neogenin in modulating BMP signaling is independent of its role in axon guidance and does not require its intracellular domain. DRAG-1/RGM physically interacts with UNC-40/neogenin, as well as the ligand and the two receptors of the BMP pathway. These results suggest a model in which the interaction between DRAG-1/RGM and UNC-40/neogenin may help the assembly of the ligand/receptor complex at the cell surface to ensure efficient BMP signaling. My work demonstrates a direct link between RGM proteins, Neogenin and BMP signaling *in vivo*, and provides a simple and genetically tractable system for further mechanistic studies of RGM and Neogenin proteins in regulation of BMP pathways.

BIOGRAPHICAL SKETCH

Chenxi Tian was born in Jin Zhou, China, a medium size city with a large population but poor economic development. She and her parents lived with her grandparents till she went to high school. Thanks to the education from the big family, especially her grandfather, she grew up loving science. She followed her heart choosing biology as a major in 2002 for her undergraduate study in a top technology and science university in China, Tsinghua University. She then came to US and became a graduate student at the Department of Molecular Biology and Genetics of Cornell University in 2006. She has been pursuing her scientific dream in the PhD program and will continue to do so.

ACKNOWLEDGMENTS

I want to thank my advisor Jun Kelly Liu for her guidance and endless help through all my PhD years, thank Ken J. Kemphues and Mariana F. Wolfner for their support and feedbacks on my projects and to make this thesis better.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

One of the fundamental questions in biology is how a single fertilized zygote divides into a large population of cells of diverse types to form a functional life unit. Cell proliferation, specification, migration and cell death are all critical cellular events in this process. Many signaling pathways and transcription factors function spatiotemporally in concert to integrate the cellular behaviors during development and homeostasis. Many human diseases, either developmental diseases or late onset diseases, are due to malfunction of these important factors. Therefore, it is beneficial for conquering human diseases to understand the molecular mechanisms underlying the regulation and function of signaling pathways and transcription factors. In the past several decades, researchers have developed multiple model organisms for studying these molecular mechanisms.

C. elegans is an excellent model organism for the study of these mechanisms. *C. elegans* is a multicellular animal that shares conserved genes with human. It has appealing advantages in several aspects in comparison to other model organisms. First, it has a short life cycle of 3 to 5 days, which is a huge advantage over mice, zebrafish, chick, and *Xenopus*, whose life cycles are much longer. This makes it ideal to perform forward genetic screens with *C. elegans*. Second, *C. elegans* has a small genome of 97Mb (*C. elegans* Sequencing Consortium, 1998), which is smaller than all the other commonly used multicellular model organisms. In the genome sequencing era, this tremendously reduces the cost and efforts of analyzing sequencing data. Third, *C. elegans* is transparent throughout its life cycle, which allows direct observation and manipulation of embryonic and postembryonic development, a task not easy to achieve in many vertebrates, especially in mice. Fourth, the cell lineage of *C.*

elegans allows us to study developmental events at single cell resolution, which is not always possible in many other multicellular animals. Last, a number of genetic manipulation techniques are available in *C. elegans*, such as powerful and effective RNA interference method, various transgenic methods to make extrachromosomal arrays, knock in and knock out genes.

In this chapter, I will start by introducing what we have learned from the development of a *C. elegans* postembryonic mesoderm lineage, the M lineage, which is my primary model system. I will then discuss a family of transcription factors, Sox, with a focus on the SoxC factors and their roles in increasing or reducing proliferation. In the third part, I will introduce the TGF-β signaling pathway and specifically two groups of TGF-β signaling modulators, RGM proteins and Neogenin.

1.1 C. elegans postembryonic mesoderm development

1.1.1 The M lineage

The postembryonic M lineage is derived from a single precursor cell, the M mesoblast (Sulston and Horvitz, 1977). It is born embryonically and divide postembryonically to give rise to all postembryonically derived non-gonadal mesodermal cells. The M cell has a reproducible division pattern. It goes through one round of dorsoventral division, one round of left-right division and two rounds of anterioposterior divisions to give rise to 16 mesoderm cells (Figure 1.1). Twelve of these cells will differentiate into the striated muscle-like body wall muscles (BWMs). Two of the 16 cells will become non-muscle macrophage-like coelomocytes (CCs), which play a role in heavy metal detoxification (Schwartz et al., 2010). The remaining two cells, M.vlpa and M.vrpa, or collectively called

M.v(l/r)pa, will divide anterioposteriorly once again. The anterior descendants will later become progenitor cells, the sex myoblasts (SMs), which will migrate to the presumptive vulva region and divide to give rise to all the non-striated sex muscles that are used for laying eggs, include type I and type II vulval muscles (VM1s and VM2s) and type I and type II uterine muscles (UM1s and UM2s). The posterior cells will differentiate into BWMs. The 14 total M lineage-derived BWMs will join the 81 embryonically derived BWMs to form the entire musculature used for locomotion.

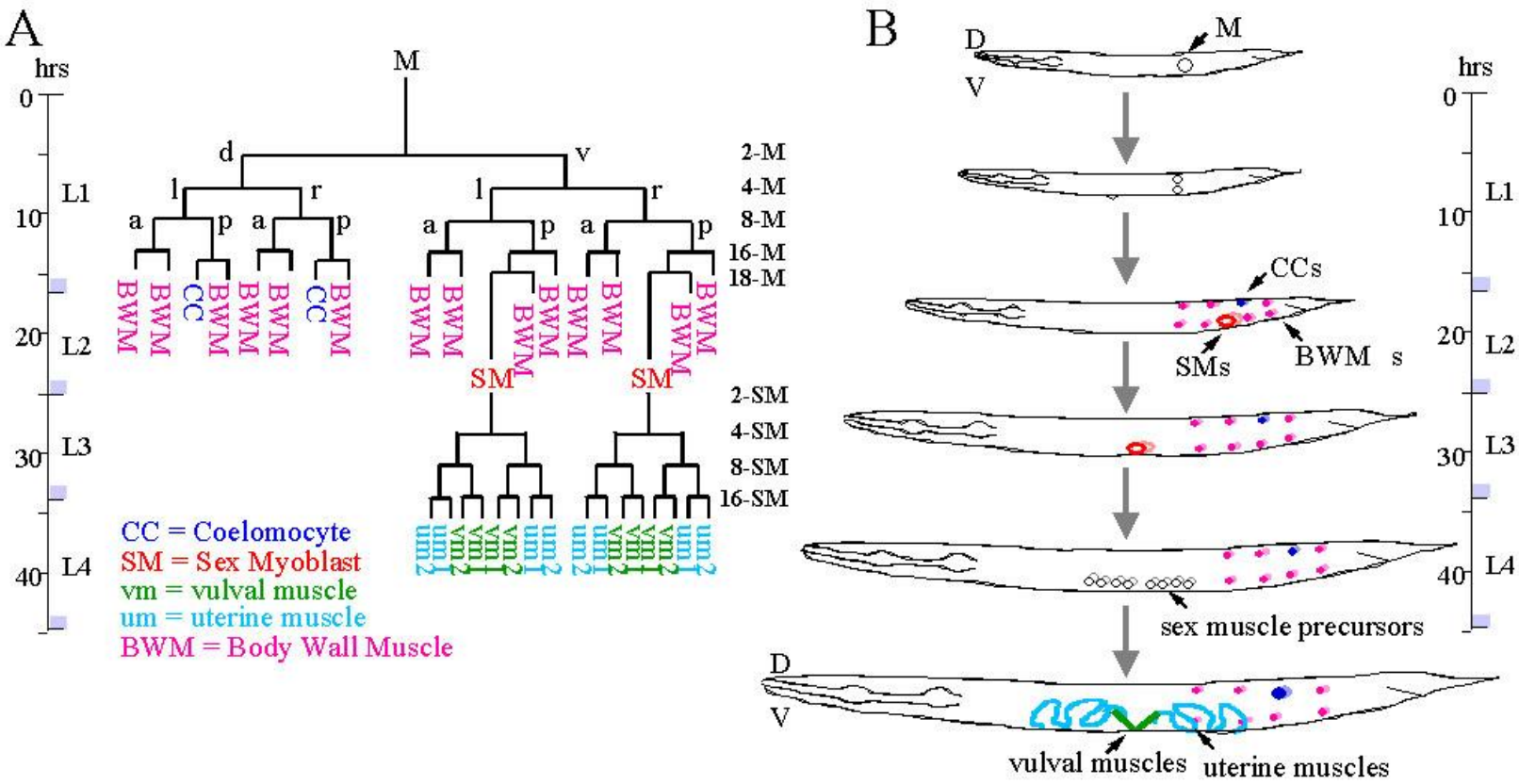
The M lineage has several advantages for studying mesodermal development. First, the M lineage is very reproducible and simple. Second, the entire lineage is not essential for the viability of *C. elegans*, allowing for various genetic and experimental manipulations. Third, a variety of cell-type-specific reporters are available to follow the M lineage development at single cell resolution (Kostas and Fire, 2002, Table 1.1).

Table 1.1: M Lineage cell-specific markers

Markers	Expression in Cell types
hlh-8::gfp	all the undifferentiated cells in the M lineage
egl-15::gfp	all VM1s
myo-3::gfp	all BMWs and VMs
Nde-box::gfp	all VMs and UMs
arg-1::gfp	VMs
CC::gfp	4 embryonic-derived CCs and 2 M-derived CCs

Figure 1.1: The *C. elegans* hermaphrodite postembryonic M lineage.

Times indicated are hours post-hatching at 25°C (modified from Sulston and Hovitz, 1977). (A) The M lineage fate map showing all the differentiated cell types. The corresponding developmental stages are indicated on both sides of the figure. (B) A schematic lateral view of developing *C. elegans* larvae with the M lineage cells highlighted. D, dorsal, V, ventral, L, left, R, right, A, anterior, P, posterior.


A
hrs
0
10
20
30
40
L1
L2
L3
L4
M
d
v
l
r
a
p
BWM
CC
SM
vm
um
2-M
4-M
8-M
16-M
18-M
2-SM
4-SM
8-SM
16-SM
CC = Coelomocyte
SM = Sex Myoblast
vm = vulval muscle
um = uterine muscle
BWM = Body Wall Muscle
B
D
V
M
CCs
SMs
BWM s
sex muscle precursors
vulval muscles
uterine muscles

The M lineage is a unique model to study mechanisms underlying several developmental processes. First of all, we can study asymmetric patterning. The M lineage exhibits both dorsoventral and anterioposterior asymmetries: the CCs are born on the dorsal side and the SMs are born on the ventral side; yet both CCs and SMs are anterior daughters of their respective mother cells. What factors regulate these asymmetries? Second, we can study the regulation of proliferation vs. differentiation. M.v(l/r)pa cells give rise to two daughter cells of entirely distinct fates: the anterior SM cells keep their proliferation ability while the posterior BWMs become fully differentiated. What regulates the decision between proliferation and differentiation? Third, we can study the different mechanisms underlying the formation of striated muscles and non-striated muscles. M.v(l/r)pa cells give rise to both striated muscles cells (BWMs) and progenitor cells for smooth muscle-like sex muscles. What are the mechanisms for the specification of each of these cell types? In the next section, I will summarize the molecular mechanisms involved in these events.

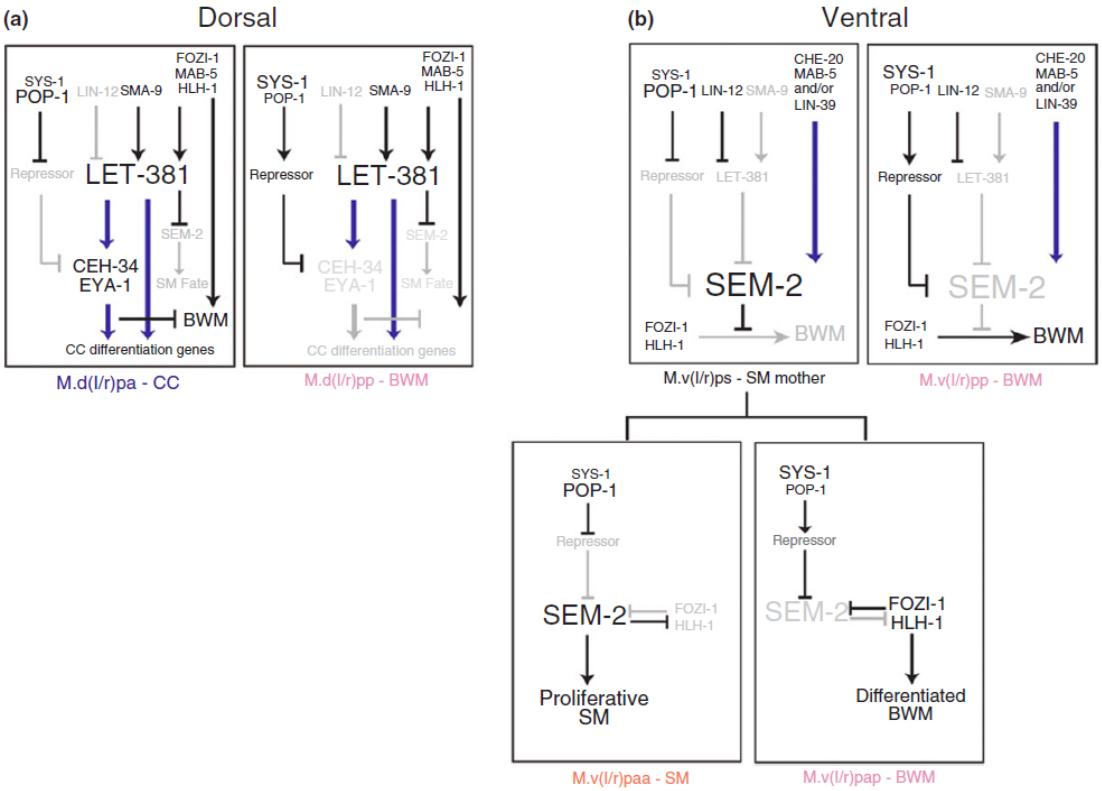
1.1.2 Mechanisms underlying the specification of M-derived cells

The specification of M-derived BWMs

Body wall muscles are derived embryonically and postembryonically (Krause and Liu, 2012). Like the embryonically derived BWMs, the specification of M-derived BWMs requires the *C. elegans* homolog of MyoD, HLH-1 (Harfe et al., 1998). The lack of *hlh-1* expression or function in the M lineage causes fate transformation from BWMs and CCs to SMs. However, only a few BWMs are affected (Harfe et al., 1998). This suggests that HLH-1 may have redundant function with other factors in specifying BWM in the M lineage. One such factor is FOZI-1. *fozi-1* encodes a zinc-finger protein with a formin homology domain (Amin et al., 2007). *fozi-1* is expressed similarly to *hlh-1* in the M lineage before the BWMs are differentiated. A *fozi-1* mutant exhibits similar fate transformation as *hlh-1* mutant: a few BWMs and all M-derived CCs are transformed to SMs. Animals that are double mutant for *hlh-1* and *fozi-1* lose most, if not all, of the M-derived BWMs. Together with the fact that *fozi-1* and *hlh-1* do not regulate each other's expression, this suggests that they function redundantly in specifying BWMs (Amin et al., 2007). The specification of M-derived BWMs also requires MAB-5, an Antennapedia homolog. *mab-5* mutants exhibit similar fate transformations as *hlh-1* and *fozi-1* mutants. Double and triple mutant analysis indicates that *mab-5* functions together with *fozi-1*, and both function redundantly with *hlh-1* (Amin et al., 2007; Figure 1.2).

Figure 1.2: Genetic regulatory network for patterning the M lineage.

Thick blue lines indicate direct regulation. Black lines represent regulation based on genetic evidence, which do not distinguish direct from indirect regulation. Grey lines and text indicate lack of expression in the corresponding cells. (A) The dorsal M lineage at the 16-M stage showing the posterior two cells. (B) The ventral M lineage at the 16-M stage showing the posterior two cells and the two descendants of the anterior cell after its division.


(a)
Dorsal
SYS-1
POP-1
LIN-12
SMA-9
FOZI-1
MAB-5
HLH-1
Repressor
LET-381
SEM-2
CEH-34
EYA-1
SM Fate
BWM
CC differentiation genes
M.d(l/r)pa - CC
M.d(l/r)pp - BWM
(b)
Ventral
CHE-20
MAB-5
and/or
LIN-39
SEM-2
FOZI-1
HLH-1
BWM
M.v(l/r)ps - SM mother
M.v(l/r)pp - BWM
Proliferative
SM
Differentiated
BWM
M.v(l/r)paa - SM
M.v(l/r)pap - BWM

M lineage expression of *hlh-1* is activated by the HMX homeodomain protein MLS-2 (Jiang et al., 2005). *mls-2* expression is likely directly activated by the PBC homeodomain protein CEH-20 independent of Hox factors (Jiang et al., 2005;Jiang et al., 2008). The M lineage expression of *fozi-1* is also regulated by *ceh-20*. However, this regulation is not mediated by *mls-2* (Amin et al., 2007).

The specification of M-derived CCs

The M-derived CCs are specified by the transient expression of the SIX homeodomain protein CEH-34 and its cofactor EYA-1 (Amin et al., 2009; Figure 1.2). *ceh-34* expression in the M lineage is directly activated by FoxF/C factor LET-381, which functions in a feed-forward manner with CEH-34 and EYA-1 to directly activate genes that are involved in CC differentiation and function (Amin et al., 2010). HLH-1 and FOZI-1 factors that specify M derived-BWMs are also required to specify M derived-CCs by positively regulating *let-381* expression (Amin et al., 2010). The expression of *ceh-34* and *eya-1* is turned on in the M.d(l/r)pa cells due to the action of several signaling pathways. The patterning along the dorsal-ventral axis requires the Schnurri homolog SMA-9, which functions to antagonize a BMP-like pathway to pattern the dorsal M lineage, as well as the LIN-12/Notch pathway that patterns the ventral M lineage (Foehr et al., 2006;Foehr and Liu, 2008). The expression of *ceh-34* and *eya-1* in the anterior cells (M.d(l/r)p**a**) requires the Wnt/β-catenin asymmetry pathway along the anterioposterior axis (Amin et al., 2009;Amin et al., 2010).

The specification of SMs

The SMs are specified by the SoxC factor SEM-2 (Tian et al., 2011). Please refer to section 1.2 for a detailed introduction to Sox factors. *sem-2* is expressed in the SM mother cells, SMs and their proliferating descendants. *sem-2* is both required and sufficient for the proliferative SM fate over the differentiated BWM fate. *sem-2* expression is directly activated by the Hox factors MAB-5 and LIN-39 and their cofactor CEH-20/PBC (Figure 1.2). The same D-V and A-P cues involved in CC specification are also involved in restricting *sem-2* expression in the M lineage (Tian et al., 2011). Notably, the D-V cues likely converge on LET-381/FoxF/C, which represses *sem-2* expression on the dorsal side of the M lineage (Tian et al., 2011; Figure 1.2). My unpublished result suggests that SEM-2 may be a cofactor of the Zinc finger transcription factor SEM-4/SALL in specifying the SM fate (See Appendix 1). *sem-2* and *sem-4* mutants share the same SM to BWM fate transformation. The onset of *sem-4* expression is one division earlier than *sem-2* in the M lineage. However, *sem-2* and *sem-4* do not regulate each other's expression in the M lineage (See Appendix 1). It will be of great interest to identify the downstream effectors of *sem-2* and *sem-4* for specifying a proliferative blast cell fate.

The specification of VMs (vulval muscles) and UMs (uterine muscles)

The VMs and UMs are non-striated smooth muscle-like cells differentiated from the SM descendants. There is very little known about factors that regulate the VM and UM fates. The T-box factor MLS-1 is both required and sufficient for specifying the uterine muscle fate (Kostas and Fire, 2002). The Meis homeodomain protein UNC-62 is required to properly specify UMs and VM2s

(Jiang et al., 2009). The Liu lab also discovered a role of the LIN-12/Notch signaling pathway in properly specifying VM2s (J. Hale and J. Liu, personal communication). In general, specification of the smooth muscles is less well studied than that of striated muscles. Therefore, characterizing key factors in the specification of VMs and UMs may provide insight into mechanisms involved in smooth muscle specification.

1.2 SoxC family members and their roles in development and disease

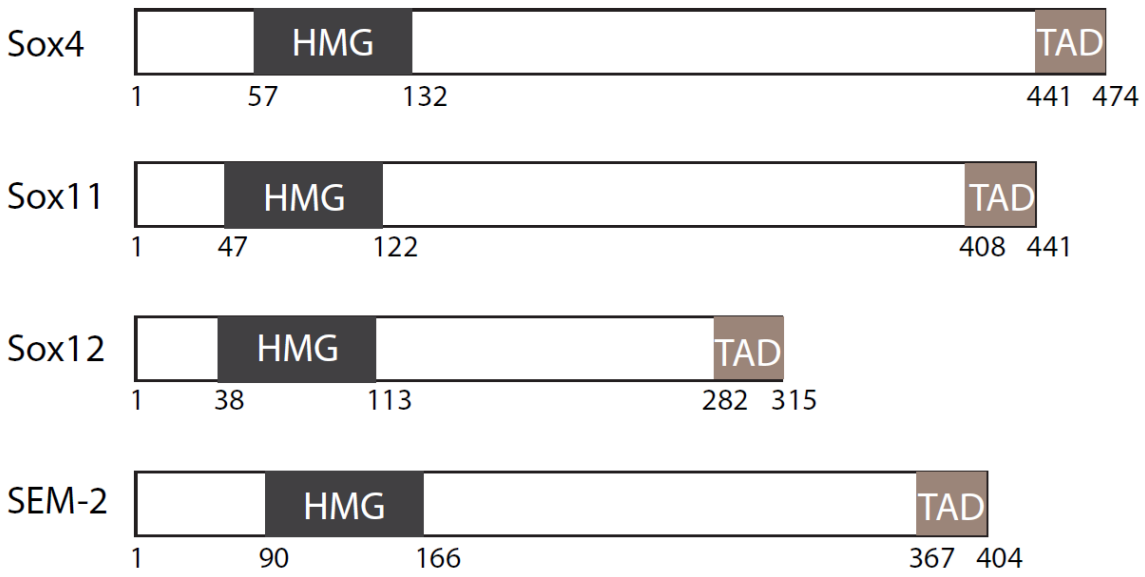
1.2.1 SoxC molecular features

Sox family members are found throughout the animal kingdom. They are characterized by the presence of a DNA-binding HMG (high mobility group) domain (Bowles et al., 2000). Sox family members are divided into groups A-J based on homologies in the HMG domain (Bowles et al., 2000). The SoxC family of Sry-related HMG box transcription factors is constituted of Sox4, Sox11 and Sox12 in vertebrates. There is only a single member of the SoxC family in *C. elegans* and *Drosophila* (Bowles et al., 2000). SoxC transcription factors are conventional transcription factors: they have a SRY-related HMG box (Sox) DNA-binding domain located in the middle of the proteins and a transactivation domain (TAD) located at the C terminus (Dy et al., 2008) (Figure 1.3). The HMG box is 84% identical at the amino acid level across all vertebrate SoxC proteins (Dy et al., 2008). The Sox4 HMG box has high affinity to the minor groove of the heptamer motif AACAAAG (van de Wetering et al., 1993). Besides the high affinity primary binding sites, Sox4 (and possibly others) is also capable of binding to secondary sites that deviate in sequences from the primary site. The crystal structure of Sox4 HMG domain binding to DNA indicates that Sox4 is capable of binding to primary and secondary sites by the concerted re-adjustment of the H-bonding pattern of two amino acid side chains (Jauch et al., 2012). Sox4 HMG box binds to DNA more efficiently than Sox12 in electrophoretic mobility shift assays (EMSA) and Sox11 has very low affinity to DNA *in vitro* (Dy et al., 2008).

The transactivation domain (TAD) is 67% identical at the amino acid level across all the SoxC members. The TAD is highly enriched in proline, serine, and acidic residues, which are commonly seen in TADs. Secondary structure predictions suggest the formation of a 20-residue alpha helix that is intact in Sox11 and disrupted in Sox4 and Sox12 (Dy et al., 2008).

Figure 1.3: Structural features of human SoxC proteins and SEM-2.

HMG: high mobility group box DNA-binding domain; TAD, transactivation domain. Numbers show the beginning and ending of each domain (Penzo-Mendez, 2010, for human SoxC proteins).


Sox4
HMG
TAD
1
57
132
441
474
Sox11
HMG
TAD
1
47
122
408
441
Sox12
HMG
TAD
1
38
113
282
315
SEM-2
HMG
TAD
1
90
166
367
404

SoxC proteins, like the SoxB1 and SoxE proteins, strongly synergize with the POU domain protein Brn1 and Brn2 to activate downstream genes that contain adjacent Sox and POU binding sites, including *Fgf4* and *Nestin* (Dy et al., 2008;Kuhlbrodt et al., 1998;Tanaka et al., 2004). Sox4 also binds to plakoglobin in response to WNT3A to inhibit transcriptional activity for some Wnt and Sox4 downstream genes (Lai et al., 2011). In addition to its role in transcriptional regulation, Sox4 was also found to regulate p53 stability and activity in lung non-small cell carcinoma cells (Pan et al., 2009a). Sox4 regulation of p53 stability is dependent on protein-protein interaction that occurs between the Sox4 DNA-binding domain and the p53 DNA-binding domain and regulatory domain (Pan et al., 2009a).

1.2.2 SoxC expression and mutant phenotypes

The expression patterns of the *SoxC* genes are largely overlapping in mice. The RNAs for all three genes were found in many regions of the central and peripheral nervous system, including the brain, neural tube, retina, olfactory epithelium, cochlear epithelium and dorsal root ganglia (Hargrave et al., 1997;Dy et al., 2008). The three genes are also co-expressed in the undifferentiated mesenchyme, such as skeletal primordia and genital tubercle, and in developing organs, such as endocaridial cushions, lung, gut, kidney and pancreas (Sock et al., 2004;Dy et al., 2008). The three genes also have gene-specific regions of expression. For example, the eyelid primordium and palatal shelf mesenchyme express *Sox11* and *Sox12*, but not *Sox4*; the developing heart endocardial cushions express *Sox4* and *Sox12*, but not *Sox11*; and the developing teeth, spleen, thymus and hair follicles predominantly express *Sox4* (Dy et al., 2008). The expression patterns of the *SoxC* genes suggest that they may play important

roles in development and this is consistent with mouse knockout mutant phenotypes. *Sox4* null embryos die at embryonic day 14 with heart outflow tract defects and incomplete ventricular septation caused by impaired development of endocardial ridges (Schilham et al., 1996). *Sox11* null mice die at birth from heart malfunctions and the phenotypes are similar to, but less severe than, those of *Sox4* null mutants (Sock et al., 2004). *Sox11* and *Sox4* null mice also display many other developmental defects, such as various craniofacial and skeletal malformations, asplenia, and hypoplasia of the lung, stomach, and pancreas (Sock et al., 2004). *Sox12* null mice have no obvious defects and have a normal life span and fertility (Hoser et al., 2008). Interestingly, although the *SoxC* genes are highly expressed in central and peripheral nervous system, no nervous system phenotypes were observed in signle mutant, suggesting that they may function redundantly there.

1.2.3 SoxC in promoting differentiation and proliferation

Overexpression and knockdown experiments of SoxC proteins suggest that they have neuronal functions. Neural tube electroporations in chicken embryos indicated that *Sox4* and *Sox11* promote neuronal gene upregulation in precursors of CNS neurons (Bergsland et al., 2006), and overexpression of *Sox4* and *Sox11* in the mouse has provided evidence for a negative influence of *Sox4* and *Sox11* on the differentiation of CNS glia (Hoser et al., 2007;Potzner et al., 2007). Recent studies showed that *Sox11* is enriched on the promoters of neuron-specific genes in embryonic stem cell-derived neurons, suggesting that SoxC proteins may be key regulators of neuronal differentiation (Bergsland et al., 2011). Consistent with this notion, *Sox4* and *Sox11* exhibit overlapping expression in the hippocampal neurogenic lineage. Overexpression of either of

them is sufficient to induce neuronal marker expression in adult neural stem cells (NSCs), while loss of *Sox4* and *Sox11* results in loss of expression of neuron-specific proteins *in vivo* and *in vitro* (Mu et al., 2012). Besides their roles in neuronal cell differentiation, SoxC factors also play roles in promoting proliferation and survival. Cell survival is reduced in the developing spinal cord with combined loss of Sox4 and Sox11 proteins, while cell proliferation is not affected (Thein et al., 2010). *Sox11* regulates survival and axonal growth of embryonic sensory neurons (Lin et al., 2011). During development of the sympathetic nervous system, *Sox11* expression turns on first and is required for proliferation in early sympathetic ganglia, whereas *Sox4* is prevalent later and required for their survival at later stages. *Sox4* and *Sox11* do not appear to function in promoting differentiation as in the central nervous system (Potzner et al., 2010). Therefore, depending on the cellular contexts, SoxC factors can promote differentiation, proliferation or survival in the nervous system.

SoxC proteins also function similarly in other cell lineages. Sox4 is required for B lymphocyte differentiation as *Sox4* null hematopoietic cells grafted into wild-type mice remain undifferentiated at the pro-B cell stage (Schilham et al., 1996). *Sox4*(+/-) mice have inhibition of proliferation, differentiation and mineralization in primary calvarial osteoblasts (Nissen-Meyer et al., 2007). Furthermore, *Sox4* null pancreatic explants displayed severely reduced numbers of endocrine insulin-producing β cells (Wilson et al., 2005). Similarly, knocking down *Sox4b* in zebrafish embryos results in loss of glucagon-producing alpha cells (Mavropoulos et al., 2005). Study of compound mutants of *SoxC* genes suggest that they are redundantly required to control survival of neural and mesenchymal progenitors and function at least in part by activating Teads,

encoding a transcriptional mediator of the Hippo pathway (Bhattaram et al., 2010).

Besides the versatile roles of SoxC factors in mouse development, SOXC factors have been implicated in human diseases in recent years. A genome wide association study identified *SOX4* as a novel gene affecting bone mineral density and fracture risk (Duncan et al., 2011). Further experiments will be required to elucidate the mechanisms.

More studies have implicated *SoxC* genes in cancer. *SOX4* or *SOX11* expression is upregulated in medulloblastomas (Lee et al., 2002), gliomas(Weigle et al., 2005), non-B cell lymphomas (Wang et al., 2008a), epithelial ovarian tumors (Brennan et al., 2009), bladder (Aaboe et al., 2006), colon(Andersen et al., 2009), prostate (Liu et al., 2006), and non-small cell lung carcinomas (Medina et al., 2009). These genes have been reported to be indicators of either good or poor prognosis in cancer patients. However, the roles of the *SOXC* genes in tumors are not fully understood, for example, the role of *SOX4* in apoptosis is controversial. *SOX4* knockdown resulted in apoptosis in ACC3 adenoid cystic carcinoma cells (Pramoonjago et al., 2006), while *SOX4* overexpression was shown to promote cell cycle arrest and apoptosis in HCT116 colon carcinoma cells (Pan et al., 2009b). Overexpression of *Sox11* prevented tumorigenesis of NSCL61s, a mouse glioma cell line, by inducing their neuronal differentiation (Hide et al., 2009). In consistent with that, higher level of *SOX4* and *SOX11* expression indicates higher survival rates in medulloblastomas and other tumor types (de Bont et al., 2008). Besides the SoxC genes' possible roles in cell survival and differentiation, *Sox4* is targeted by microRNA miR-335 to inhibit metastatic cell invasion by acting at least in part through the *Sox4* target

gene Tenascin, a known extracelluar matrix protein implicated in cell migration (Tavazoie et al., 2008).

In summary, SOXC proteins have possible but as yet unclear roles in cell survival, apoptosis, differentiation and metastatic invasion, and the different roles may depend on the context of different cell lineages or tumor types. Further studies are required to elucidate the mechanisms of the involvement of SoxC factors in development and diseases, especially cancer.

1.3 The TGF-β signaling pathway

1.3.1 The core TGF-β pathway

The transforming growth factorβ (TGFβ) superfamily of extracellular signaling molecules regulates a wide variety of cellular processes, including cell fate specification, cell proliferation, cell migration, and cell death throughout all developmental stages of multi-cellular organisms (reviewed in Massague et al., 2000b). Malfunction of the pathway causes many somatic and hereditary disorders in humans, including cardiovascular diseases and various types of cancers (Massague et al., 2000a;Siegel and Massague, 2003;Gordon and Blobe, 2008;Massague, 2008). The TGFβ family of cytokines includes two sub-groups defined by sequence similarity and function: the TGFβ/activin/Nodal subfamily and the BMP (Bone morphogenetic protein)/GDF (Growth and differentiation factor)/MIS (Muellerian inhibiting substance) subfamily (Shi and Massague, 2003). The TGFβ family ligands are generated with a prodomain proceeding the mature domain. The prodomain is cleaved by furin protease in the secretory

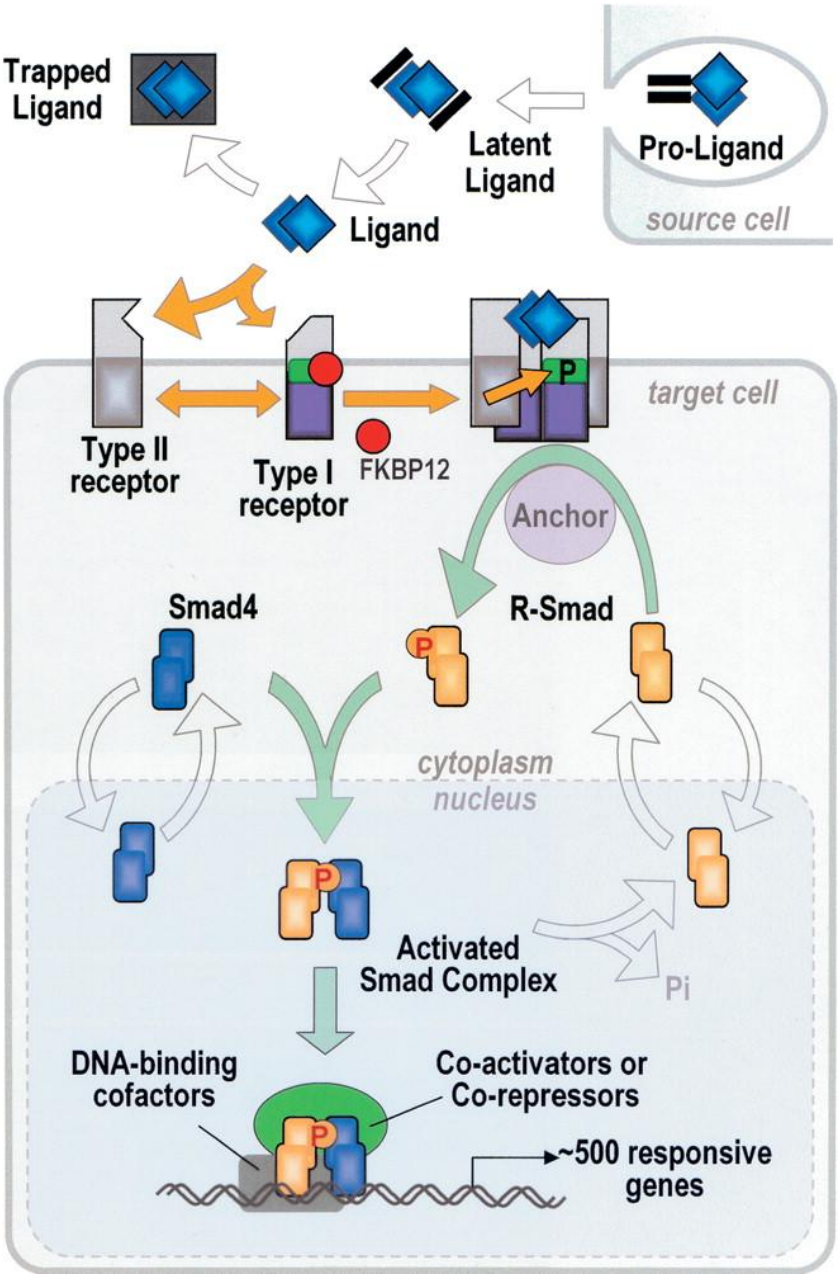
pathway but remains non-covalently bound to the active domain (Miyazono et al. 1988; Wakefield et al. 1988). Activation of the ligands requires removal of the prodomains (Figure 1.4). The active forms of ligands of different subfamilies share structural similarities, including a dimer stabilized by hydrophobic interactions and an inter-subunit disulfide bond (Shi and Massague, 2003).

The receptors for TGFβ cytokines belong to the receptor serine/threonine kinase family. TGFβ receptors fall into two classes: type I and type II (Attisano et al., 1994;Massague, 1992;Massague, 1998). TGF-β and BMP subfamily ligands utilized different sets of type I and type II receptors. Both types of receptors are type I transmembrane proteins, comprised of an extracellular domain, a transmembrane region, and a C-terminal serine/threonine kinase domain. The extracellular domains of type I and type II receptors are capable of binding to the ligands. Two distinct modes of ligand-receptor interaction exist. BMP ligands exhibit a high affinity for the extracellular domains of the type I receptors and a low affinity for the type II receptors. The ligand binds to type I receptor first and the ligand-type I receptor complex subsequently binds to type II receptors. In contrast to the BMPs, TGFβ and Activin display a high affinity for the type II receptors and do not interact with the type I receptors. Therefore, the TGFβ ligand binds tightly to type II receptor first; this binding allows subsequent incorporation of the type I receptor. The type I, but not type II, receptors contain a characteristic SGSGSG sequence, also known as the GS domain, which is located N-terminal to the kinase domain. Once the ligand binding brings the type I and type II receptors to close proximity, constitutively active type II receptor phosphorylates type I receptor at its GS domain (Figure 1.4).

The activated type I receptor is poised to relay the TGF-β signal through the Smad proteins, which transduce the signal from the cytoplasm to the nucleus (Massague, 1998;Massague and Chen, 2000;Massague and Wotton, 2000b;Massague et al., 2005a; Figure 1.4). There are two classes of Smad proteins required for signal transduction: receptor regulated Smads (R-Smads) and collaborating Smads (co-Smads). R-Smads are directly phosphorylated and activated by the type I receptors. The TGFβ subfamily uses Smad2 and Smad3, while the BMP subfamily transduces through Smads 1, 5, and 8.

Figure 1.4: overview of the TGF-β signaling pathway

TGFβ family members are dimeric ligands that remain in latent form by binding to their propeptides, or in trapped forms by binding to different ligand traps. Only when released from these inactive states, the ligands bind to two pairs of type I and type II receptor serine/threonine kinases, forming a hetero-tetrameric receptor complex. The cytoplasmic region of the type I receptor contains a canonical protein kinase domain (purple) C terminal to a regulatory region or GS domain (green). Once in close proximity, type II receptor phosphorylates the GS domain of type I receptor, which then activate R-Smads by phosphorylating the SXS motif at the C terminus of R-Smads. The phosphorylation at the C terminus of R-Smads creates a pS-x-pS motif (red P), allowing the accumulation of R-Smads in the nucleus and the recognition of this motif by Smad4. The Smads complex then recruits DNA-binding cofactors and transcription co-activators and co-repressors to confer transcriptional control. Modified from Massagué et al. 2005b.


Trapped Ligand
Latent Ligand
Pro-Ligand
source cell
Ligand
target cell
Type II receptor
Type I receptor
FKBP12
P
Anchor
Smad4
R-Smad
cytoplasm
nucleus
Activated Smad Complex
Pi
DNA-binding cofactors
Co-activators or Co-repressors
~500 responsive genes

The Smad proteins share similar structural features, including the N-terminal MH1 and the C-terminal MH2 domains. The MH1 domains exhibit DNA sequence specific binding activities, and inhibit MH2 domain function. The MH2 domain is required for Smad-receptor interaction, Smad-Smad interaction and Smad-Smad DNA-binding cofactor interactions. It also directly interacts with the cytoplasmic retention proteins and the nuclear pore compex for nucleocytoplasmic shuttling. The R-Smads, but not the Co-Smads, exhibit a Ser-X-Ser (SXS) motif at the C terminus (Figure 1.4). This SXS motif is phosphorylated by the activated type I receptor. The phosphorylated C terminus provides a binding site for Smad4, the Co-Smad, which is an essential component in the assembly of target-specific transcriptional complexes. R-Smad translocation does not require the presence of co-Smads. However, co-Smads require the presence of activated R-Smads to travel into the nucleus (Attisano and Wrana, 2000;Massague et al., 2005a). The resulting Smad complex incorporates different DNA-binding cofactors that confer target gene selectivity and affect the recruitment of either transcriptional co-activators or co-repressors (Figure 1.4).

The MH1 domain (approximately 130 amino acids) and the MH2 domain (approximately 200 amino acids) are separated by a linker domain, which is not conserved in sequence or length (Massague, 1998). However, the linker region contains multiple phosphorylation sites that have critical functions. When directed to the nucleus by TGF-β or BMP signals, the linker region is phosphorylated by MAP kinases and cell division cyclin-dependent kinases (CDKs) in response to mitogens and stresses to constrain TGF-β and BMP signaling (Kretzschmar et al., 1997;Kretzschmar et al., 1999;Matsuura et al., 2010). Smad proteins can also be phosphorylated by cyclin-dependent kinase 8/9 (CDK8/9) and glycogen synthase

kinase-3 (GSK3), and these phosphorylations are required for transcriptional action and subsequently required for Smad destruction (Aragon et al., 2011).

1.3.2 The modulation of the TGF-β pathways

The TGF-β signaling pathway can be modulated at different levels, including receptor activation outside the cell, Smad access to the receptors in the cytoplasm, Smad nucleocytoplasmic shuttling, and Smad-dependent transcription in the nucleus (Shi and Massague, 2003). I will mainly focus on modulations of the TGF-β pathway that take place outside the cell. Please refer to these excellent reviews for modulation of TGF-β signaling at other levels (Massague, 1998;Massague and Chen, 2000;Massague and Wotton, 2000a;Shi and Massague, 2003; Chen YG, 2009).

At the level of receptor activation, extracellular secreted proteins, such as Noggin, Chordin/Sog, Gremlin, Follistatin, and DAN/Cerberus, have been shown to function as ligand binding traps (Balemans and Van Hul, 2002;Shi and Massague, 2003) (Figure 1.4). Noggin is the best known BMP antagonist. Noggin sequesters the ligands and blocks the ligands from binding to the type I and II receptors (Groppe et al., 2002). In Drosophila, Twisted gastrulation (Tsg) forms a complex with the Chordin homolog Short gastrulation (Sog). This complex sequesters BMPs to attenuate BMP signaling locally but enhances long range BMP signaling by transporting BMPs through tissues (O'Connor et al., 2006). Crossveinless 2, another Drosophila homolog of Chordin, can activate or inhibit BMP signaling through directly binding to the BMP ligands, and the different responses are determined by different BMP types and concentrations (Ambrosio et al., 2008;Serpe et al., 2008;Zhang et al., 2008).

The extracellular matrix (ECM) proteins have also been shown to modulate TGF-β signaling. β–glycan, also known as the TGF-β type III receptor, mediates TGF-β2 binding to type II TGF-β receptor and also increases the affinity of Inhibin to Activin and type II BMP receptors (del Re et al., 2004;Lewis et al., 2000;Wiater et al., 2006). A glycosylphosphatidylinositol (GPI)-anchored proteoglycan family member in Drosophila, Dally, positively regulates the distribution of Dpp and Wg morphogen (Grisaru et al., 2001;Han et al., 2005;Takeo et al., 2005). Cripto, an epidermal growth factor-CFC (Cripto/FRL-1/Cryptic) motif-containing GPI-anchored membrane protein, acts as a co-receptor to mediate the binding of Nodal and GDF1 to Activin receptors. Type IV collagens were found to modulate Dpp signaling by affecting the formation of Dpp gradient in Drosophila (Wang et al., 2008b). RGM (Repulsive Guidance Molecule) proteins are co-receptors for BMP signaling (Babitt et al., 2005;Babitt et al., 2006;Xia et al., 2007). A RGM interacting protein Neogenin has been implicated in modulating BMP signaling (Zhou et al., 2010;Hagihara et al., 2011). RGMs and Neogenin will be the focus for sections 1.4 and 1.5, respectively.

1.3.3 The TGF-β pathways in C. elegans

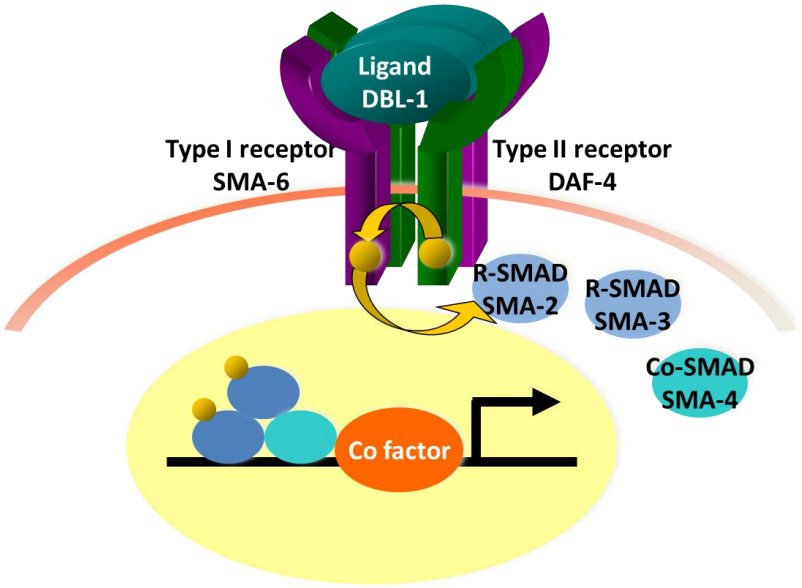
There are two BMP related TGF-β signaling pathways in *C. elegans*: the Sma/Mab pathway, which regulates body size (Savage-Dunn et al., 2003;Suzuki et al., 1999;Suzuki et al., 1999), male tail patterning (Savage-Dunn et al., 2003;Suzuki et al., 1999;Suzuki et al., 1999), and M lineage patterning (Foehr et al., 2006), and the dauer pathway, which controls formation of dauer larvae in response to stresses (Ren et al., 1996). These two pathways share the same type II receptor DAF-4 (Estevez et al., 1993). All other components are pathway-specific. The Sma/Mab TGF-β pathway consists of the ligand DBL-1, the type I

receptor SMA-6, the R-Smads SMA-2 and SMA-3, and the co-Smad SMA-4 (Patterson and Padgett, 2000;Savage-Dunn, 2005; Figure 1.5). The dauer pathway consists of the ligand DAF-7, the type I receptor DAF-1, the R-Smads DAF-8 and DAF-14, and the co- or anti-Smad DAF-3 (Savage-Dunn, 2005). In addition there are two orphan TGF-β-like ligands, UNC-129 and TIG-2. UNC-129 functions in the axon guidance pathway and does not seem to signal through any of the identified receptors or Smads (Colavita et al., 1998;Nash et al., 2000). The function of TIG-2 is unknown.

A few signaling components outside of the core Sma/Mab pathway have been identified. These include three modulators at the ligand-receptor level. LON-2 is a Glypican homolog in *C. elegans* that negatively regulates Sma/Mab signaling possibly by interacting with the ligand DBL-1 (Figure 1.5) (Gumienny et al., 2007). Loss-of-function mutation in *lon-2* results in long worms. Two proteins are positive modulators of the Sma/Mab signaling: CRM-1 is a cysteine-rich repeats containing protein (Fung et al., 2007); SMA-10 is a member of the LRIG (leucine rich repeats and immunoglobulin-like domains) family of transmembrane protein (Gumienny et al., 2010). LON-1 is a downstream target of the Sma/Mab pathway (Figure 1). It encodes a pathogenesis-related (PR) protein that regulates polyploidization and body length (Maduzia et al., 2002;Morita et al., 2002). LON-1 is likely also a modulator of the Sma/Mab pathway by interacting with LON-2 and CRM-1 (King Chow, personal communication). SMA-9 is a homolog of Drosophila Schnurri (Shn) protein (Liang et al., 2003). In *C. elegans*, SMA-9 is known to function as transcription co-factors to convey Sma/Mab signaling (Liang et al., 2003).

Figure 1.5: the Sma/Mab BMP-like pathway in *C. elegans*

C. elegans Sma/Mab core pathway includes the ligand DBL-1, type I receptor SMA-6, type II receptor DAF-4, R-SMads SMA-2 and SMA-3, and co-Smad SMA-4.


Ligand
DBL-1
Type I receptor
SMA-6
Type II receptor
DAF-4
R-SMAD
SMA-2
R-SMAD
SMA-3
Co-SMAD
SMA-4
Co factor

The role of the Sma/Mab pathway in patterning of the *C. elegans* M lineage was discovered in the Liu lab (Foehr et al., 2006). Mutations in *sma-9* cause a dorsal to ventral fate transformation in the M lineage, in addition to the small body size phenotype (Liang et al., 2003). The *sma-9* M lineage defects can be suppressed by mutations in the Sma/Mab pathway, including *dbl-1*, *sma-2*, *sma-3*, *sma-4*, and *sma-6*. However, these mutations themselves do not result in any M lineage defects (Foehr et al., 2006). These observations suggest that SMA-9 normally antagonizes the Sma/Mab BMP-like pathway to regulate dorsal/ventral patterning in the M lineage. Studies of additional *sma-9* suppressors have identified *drag-1/Rgm* and *unc-40/Neogenin*. They are the focus of Chapters 3 and 4.

1.4 RGM in modulating BMP signaling

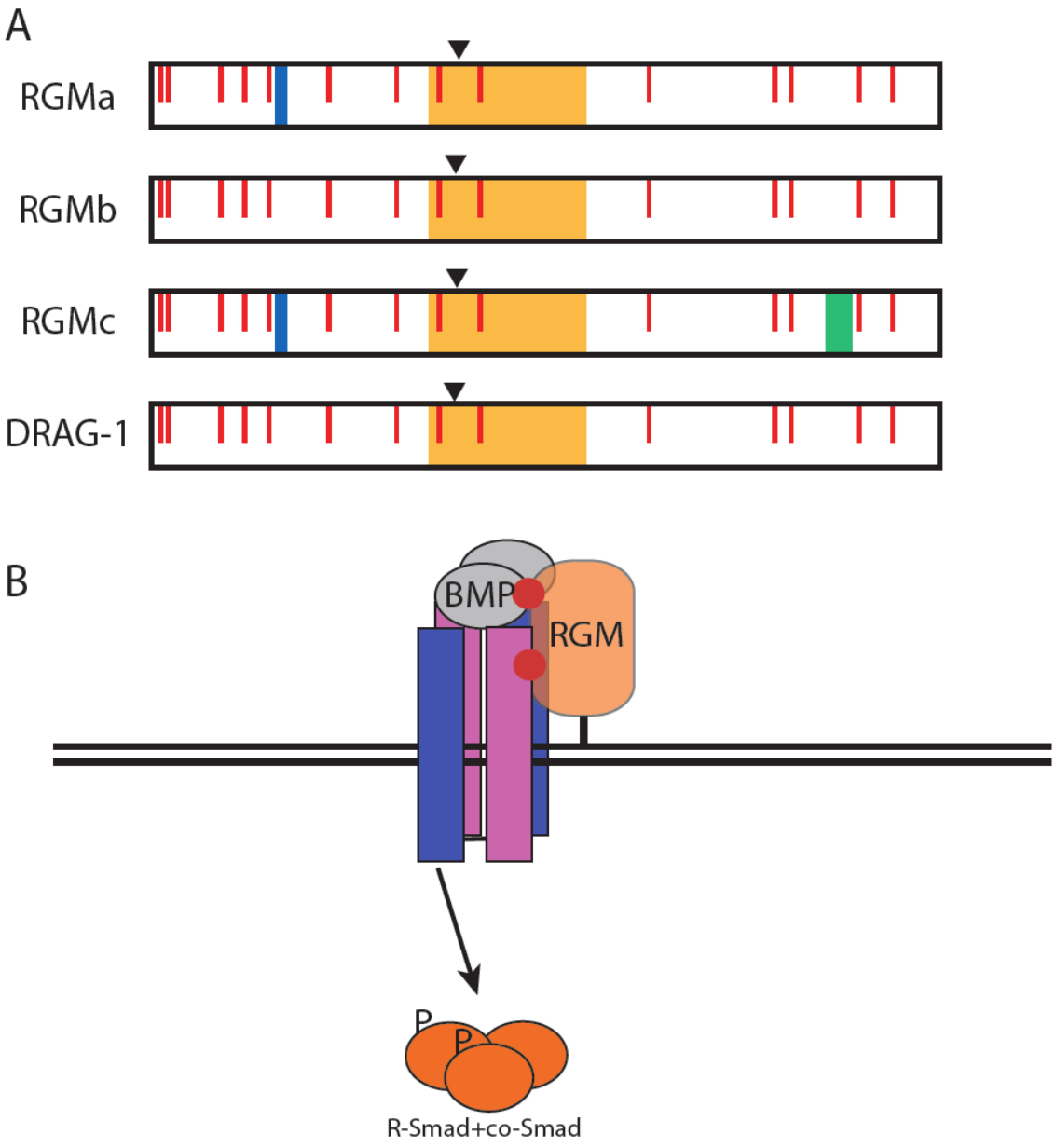
1.4.1 Introduction

RGM (repulsive guidance molecule) family members include RGMa, RGMb and RGMc/Hemojuvelin (Hjv). RGMs also exist in single copy in invertebrates, such as *C. elegans* and sea urchin (Camus and Lambert, 2007). The three RGMs share approximately 40-50% identity in primary amino acid sequences. They share similarities in predicted protein domains and overall structure, as inferred by *ab initio* molecular modeling (Severyn et al., 2009). Specifically, all RGMs have an N-terminal signal peptide, a partial von Willebrand type D domain (vWF-type D), which includes a highly conserved autoproteolysis site, and a C-terminal GPI (Glycosylphosphatidylinositol)-anchor (Figure 1.6).

RGMa and RGMc have a RGD motif (Figure 1.6). RGM proteins also contain a number of highly conserved cysteine residues. Studies of vertebrate RGM proteins demonstrate that RGMs undergo distinct biosynthetic and processing steps. All of them go through autoproteolytic cleavage to generate two fragments that are linked by disulfide bonds. RGMc also has a furin pro-protein convertase recognition site that is unique to itself. The significance of these processing events is beginning to be revealed. RGM proteins are known high affinity binding proteins for BMPs (K_D ~ 1-5 mm) (Babitt et al., 2005;Babitt et al., 2006;Samad et al., 2005; Figure 1.6). Although no role for BMPs has been identified in the actions of RGMa, interactions with BMP is mediating at least some of the biological effects of RGMc on the regulation of iron homeostasis (Babitt et al., 2006;Babitt et al., 2007), as well as the roles of RGMb in promoting neurite growth and nerve regeneration. As described in Chapter 3, the *C. elegans* RGM protein DRAG-1 is the only studied RGM in invertebrates (Tian et al., 2010). DRAG-1 exhibits structure and function similar to vertebrate RGMs. DRAG-1 protein contains the conserved partial vWF-type D domain, the autocleavage site, and all the cysteine residues (Figure 1.6). DRAG-1 is a tissue specific positive modulator of the BMP-like Sma/Mab pathway in *C. elegans* by interacting with the ligand and both type I and type II receptors of this pathway (Tian et al., 2010; Chapter 5). This section will focus on the current understanding of RGM modulation of BMP signaling.

Figure 1.6: Structure and function of RGM family members

(A) The linear maps of mature RGMa, RGMb and RGMc contain the following features: RGD motif (RGMa and RGMc, blue); partial vWF-type D domain (yellow); furin recognition and cleavage site (RGMc only, green); arrowhead, site of intra-molecular proteolytic cleavage to generate two-chain species; red vertical lines, conserved cysteine residues. The squiggle at the C-terminus of each protein represents the GPI anchor. Modified from Severyn et al., 2009. (B) RGM proteins are BMP coreceptors. RGM mediates BMP signaling by interacting with the ligand and type I and type II receptors and enhancing ligand-receptor binding. Direct interactions are indicated by red dots.


A
RGMa
RGMb
RGMc
DRAG-1
B
BMP
RGM
P
P
R-Smad+co-Smad

1.4.2 RGMa

RGMa was the first RGM family member discovered. RGMa is expressed in a gradient in the optic tectum in chick embryos and is functionally recognized as a repulsive guidance molecule (Monnier et al., 2002). During mouse development, RGMa is expressed in the central nervous system in a mostly non-overlapping fashion with RGMb. RGMa is also expressed in the developing mouse cochlea, lung, limb primordia, and gut. In the adult mouse, RGMa expression is found in the heart, brain, liver, lung, skin, kidney, testis and gut (Corradini et al., 2009). The broad expression pattern of RGMa in developing mouse embryo suggests that it likely plays important roles in development, especially central nervous system development. In fact, RGMa has been found to function in axon guidance in the developing visual system and in axon tract formation in the embryonic brain *in vivo* (Monnier et al., 2002;Matsunaga et al., 2004;Tassew et al., 2008), as well as in neural tube closure (Kee et al., 2008), neuronal differentiation (Matsunaga et al., 2006), cell survival (Koeberle et al., 2010), cell migration and adhesion (Lah and Key, 2012). Neogenin, as a receptor for RGMa, is involved in all the above processes. Recently, RGMa has been implicated in some colorectal, prostate, and breast cancers (Li et al., 2011;Li et al., 2012;Zhao et al., 2012), as well as associated with inflammation (Nohra et al., 2010;Mirakaj et al., 2011;Muramatsu et al., 2011).

RGMa was found to be a BMP coreceptor *in vitro* (Samad et al., 2005). Both overexpression and siRNA knockdown experiments showed that RGMa specifically enhances BMP signaling activity through the canonical BMP signaling pathway, but not the TGF-β signaling pathway (Babitt et al., 2005;Halbrooks et al., 2007). The enhancement depends on BMP ligands, because it was blocked

by co-transfection of Noggin, a BMP inhibitor, or BMP2 and BMP4 neutralizing antibodies (Babitt et al., 2005). Besides the dependent of the ligand, RGMa modulates BMP signaling through the classical BMP signaling pathway involving type I receptors ALK3 and ALK6 as well as Smad1, 5, and 8, because dominant negative forms of ALK3 and ALK6 both blocked the enhanced signaling and Smad 1, 5, 8 phosphorylation by RGMa (Babitt et al., 2005). Purified RGMa.Fc (the extracellular domain of RGMa fused to the Fc portion of human IgG) protein binds directly and selectively to ^{125}I-BMP2 and ^{125}I-BMP4 with a K_D of 2.4 nM and 1.4nM, respectively, but not to BMP7 or TGF-β1 ligands (Babitt et al., 2005;Halbrooks et al., 2007;Xia et al., 2007). BIAcore assay identified a similar direct binding of RGMa.Fc to BMP2 at K_D ~ 2.46nM (Babitt et al., 2005;Halbrooks et al., 2007). RGMa.Fc binds to the BMP type I receptor ALK6.Fc with or without the presence of BMP2. The binding between RGMa and ALK6 increased the binding of ^{125}I-BMP2 to ALK6 (Babitt et al., 2005). Another study showed that the addition of RGMa not only increased the binding of ^{125}I-BMP2 to the BMP type I receptor ALK3.Fc, but also increased the binding of ^{125}I-BMP2 to the type II activin receptor ActRIIA.Fc (Xia et al., 2007). A model has therefore been proposed in that the addition of RGMa changes the utilization of type II receptors from BMPRII alone to both BMPRII and ActRIIA for BMP2 or BMP4 signaling in human ovarian granulosa cells and mouse pulmonary artery smooth muscle cells (Xia et al., 2007).

The *in vivo* evidence is still lacking for RGMa functioning as a BMP co-receptor. How many of the RGMa mediated processes require its potentiation of BMP signaling? BMP signaling occurs in neurons of the adult spinal cord which expresses RGMa in mouse; however, it is unknown if RGMa is involved in BMP signaling there. In axon guidance process, it has been shown that RGMa-

neogenin signaling involves activation of the small GTPase RhoA, its downstream effector Rho kinase and protein kinase C (PKC) but is independent of the BMP signaling (Conrad et al., 2007). It is also of interest to know if neogenin, the RGMa receptor that mediates axon guidance and neuronal survival, functions together with RGMa to modulate BMP signaling.

1.4.3 RGMb

RGMb is currently less well characterized than RGMa. RGMb was first identified by using a genomic DNA-binding array to identify genes regulated by DRG11, a homeobox transcription factor expressed in embryonic dorsal root ganglion (DRG) and dorsal horn neurons (Samad et al., 2004). In the developing brain, RGMa and RGMb exhibit non-overlapping expression patterns. RGMb expression is also found in the brain, bone, heart, lung, liver, kidney, testis, ovary, uterus, epididymis and pituitary in adult mouse (Corradini et al., 2009). Unlike RGMa, RGMb does not have any detectable repulsive guidance role in embryonic and neonatal DRG neuritis (Samad et al., 2004). However, there is evidence supporting its role in the nervous system, immune system and cancer. RGMb controls aggregation and migration of neogenin-positive dentate precursor cells *in vitro* and *in vivo*, and this process might involve RGMb-neogenin interaction (Conrad et al., 2010). RGMb also promotes neurite outgrowth and peripheral nerve regeneration *in vitro* and *in vivo* (Liu et al., 2009;Ma et al., 2011). RGMb is an important negative regulator of IL-6 expression in immune cells (Xia et al., 2011). Furthermore, RGMb appears to be a negative regulator of breast cancer proliferation, adhesion, and migration *in vitro* (Li et al., 2012).

RGMb was the first RGM family member shown to be a BMP co-receptor. The reason to test its involvement in BMP signaling is due to the shared expression pattern between RGMb and the BMP type I and type II receptors in the developing mouse and *Xenopus* embryos (Babitt et al., 2005). Similar to RGMa, RGMb enhanced BMP signaling but not TGF-β signaling in LLC-PK1 (kidney epithelial cells) and in *Xenopus* embryos (Babitt et al., 2005). siRNA treatment of RGMb inhibited BMP signaling (Babitt et al., 2005;Halbrooks et al., 2007). Like RGMa, RGMb enhancement of BMP signaling is ligand-dependent and through the canonical BMP signaling pathway (Babitt et al., 2005). Also like RGMa, RGMb.Fc (the extracellular domain of RGMb fused to the Fc portion of human IgG) binds directly to ^{125}I-BMP2 (K_D ~ 1.4nM). The binding can be competed by the addition of unlabeled BMP2 and BMP4, but not BMP7 or TGF-β ligands (Babitt et al., 2005). BIAcore assay also confirmed that RGMb.Fc binds directly to BMP2 with a similar binding affinity (K_D ~ 5.43 nM; Babitt et al., 2005;Halbrooks et al., 2007). RGMb also physically associates with BMP type I receptors (ALK2, ALK3, and ALK6) and BMP type II receptors (ActRII and ActRIIB) as indicated by co-immunoprecipitation experiment in transfected HEK293 cells (Babitt et al., 2005). It has been suggested that RGMb shares the ability of RGMa to alter type II receptor utilization by BMP4 ligand from unpublished data noted in a review article (Corradini et al., 2009).

In one study, RGMb was shown to inhibit BMP signaling in C2C12 myoblasts. This inhibition by RGMb was dependent on the vWF type D domain. Because it can suppress induced BMP signaling by constitutively active Smad1, the suppression is likely mediated by a parallel pathway that is activated by RGMb (Kanomata et al., 2009). This inhibitory effect of BMP signaling seems to be an isolated event, as it has not been observed for RGMa and RGMc in any

system or for RGMb in a different system. How RGMb mediates this inhibition and its significance of it needs to be studied.

The *in vivo* relevance of RGMb as a BMP co-receptor has just started to be discovered. RGMb promotion of axon growth requires BMP signaling. The fewer and shorter neurites from cultured neonatal whole DRG explants and dissociated DRG neurons in *RGMb-/-* mutants could be rescued by the addition of BMP2. RGMb also promotes, while the BMP inhibitor Noggin inhibits, peripheral nerve regeneration *in vivo* (Ma et al., 2011). Like RGMa, RGMb also directly binds to neogenin (Conrad et al., 2010). It is not known if neogenin functions together with RGMb in mediating BMP signaling.

1.4.4 RGMc/HJV

HJV and juvenile hemochromatosis

RGMc was identified by positional cloning for the locus that is associated with juvenile hemochromatosis (JH) (Papanikolaou et al., 2004b). JH is a rare autosomal recessive disease characterized by high penetrance, early-onset systemic iron overload that affects young patients and leads to severe clinical complications typically in the first and second decade of life (Camaschella et al., 2002;De Gobbi et al., 2002;De Domenico et al., 2008a).

JH is also caused by the lack of functional hepcidin (Roetto et al., 2003). Hepcidin is a defensin-like small peptide secreted predominantly from hepatocytes (Park et al., 2001;Pigeon et al., 2001). It is essential for iron homeostasis. It acts by binding to and degrading Ferroportin, a rate-limiting sole iron exporter (De Domenico et al., 2008b). Hepcidin expression is upregulated by

high iron level in the body, therefore it is a feedback regulator of iron absorption (Kautz et al., 2008a;Pigeon et al., 2001).

HJV is a key regulator of hepcidin expression

HJV is both required and sufficient for hepcidin expression. JH patients with HJV mutations exhibit a reduction of hepatic hepcidin expression that results in severe iron accumulation in liver, heart, and pancreas (Papanikolaou et al., 2004a). *Hjv-/-* mice exhibit a similar low hepcidin mRNA expression and iron deposition in liver, pancreas, and heart (Huang et al., 2005). Inhibition of HJV function by a soluble HJV.Fc reduces hepcidin expression (220 Babitt,J.L. 2007). Transfection with *HJV* cDNA into hepatoma-derived cells increases hepcidin mRNA expression and hepcidin promoter activity in a luciferase assay (Babitt et al., 2006). Introducing Hjv into the hepatocytes in *Hjv-/-* mice using adeno-associated virus 2/8 as a vector is sufficient to completely restore the decreased hepatic hepcidin expression and lower the high serum iron level to the wild type level (Zhang et al., 2010).

Consistent with HJV regulating hepcidin expression in the liver, *HJV* is expressed predominantly in the liver, skeletal muscle, and heart (Niederkofler et al., 2004;Papanikolaou et al., 2004a;Rodriguez et al., 2007a;Samad et al., 2004). The functional significance for skeletal muscle and heart-expressed HJV is currently not fully understood. Skeletal muscle-expressed HJV has been suggested to be a reservoir for soluble HJV (Zhang, 2010; see section on HJV post-translational modification).

BMP regulates hepcidin expression

BMP signaling has been shown to play a role in iron metabolism in liver by regulating Hepcidin expression. *In vitro* studies showed that multiple BMP cytokines, including BMP-2, 4, 5, 6, 7, 9 are able to increase hepcidin expression through phosphorylated Smad1, 5, and 8 in hepatocytes (Babitt et al., 2006;Truksa et al., 2006;Andriopoulos et al., 2009a). Liver-specific disruption of Smad4, the co-Smad, affected iron homeostasis by decreasing Hepcidin expression (Wang et al., 2005).

HJV is a BMP co-receptor

The underlying mechanism of HJV regulating hepcidin expression is through BMP signaling. Like other RGM family members, HJV functions as a co-receptor in BMP signaling. Transfection of HJV into hepatocytes enhances BMP, but not TGF-β signaling activity (Babitt et al., 2005;Babitt et al., 2006;Halbrooks et al., 2007;Xia et al., 2008). si-RNA knockdown of HJV reduced BMP2 signaling (Babitt et al., 2005;Halbrooks et al., 2007). Similar to RGMa and RGMb, HJV's role in potentiating BMP signaling is dependent on the BMP ligand, because HJV signal induction is abolished by the BMP inhibitor Noggin, and by BMP2 and BMP4 neutralizing antibodies (Babitt et al., 2006). HJV signals via type I and II receptors and the Smads complex. Inhibition of type I receptors (ALK2, ALK3, and ALK6) and type II receptors (BMPRII and ACTRII) by si-RNA led to decease in HJV-mediated BMP signaling (Xia et al., 2008). Like RGMa and RGMb, purified HJV.Fc directly binds to ^{125}I-BMP2 and ^{125}I-BMP4, and this binding can be competed away by excess unlabeled BMP2 and BMP4, but not by BMP7 or TGFβ1. HJV.Fc also directly binds to BMP6 (Andriopoulos et al., 2009a). The binding affinity of HJV to BMP2 has been estimated by two assays with distinct

values: one BIAcore assay showed that dimeric HJV.Fc binds to BMP2 with an affinity similar to RGMa and RGMb (K_D ~ 4.22nM; Babitt et al., 2005;Halbrooks et al., 2007); another BIAcore assay with purified untagged monomeric HJV and BMP2 estimated the binding K_D to be around 510nM (Yang et al., 2008b). This suggests that dimeric form of HJV has a stronger affinity to BMP than the monomeric form. HJV has a different binding selectivity to BMP ligands than the other RGM family members. Soluble RGMb.Fc and HJV.Fc were assayed for the ability of inhibiting selected BMP signaling activity. HJV.Fc is a stronger inhibitor to BMP-6 than RGMb.Fc, while RGMb.Fc is a stronger inhibitor than HJV.Fc in inhibiting BMP-2 and BMP-4 signaling (Andriopoulos et al., 2009a). In addition to binding to the ligand, HJV can form a complex with type I receptor ALK6 in the presence of BMP2 in a co-immunoprecipitation assay (Babitt et al., 2006).

Consistent with the *in vitro* data, *Hjv-/-* mice have reduced hepatic hepcidin expression and a decreased level of liver pSmad, an indication of low BMP signaling activity. Hepatic expression of Hjv in *Hjv-/-* mice liver increased hepcidin expression and pSmad level (Zhang et al., 2010).

BMP6 is involved in iron-regulated hepatic hepcidin expression

BMP6 mRNA expression has been found to correlate to the extent of body iron load (Kautz et al., 2008b). *Bmp-6-/-* mice exhibit a severe loss of hepcidin expression and massive iron overload (Andriopoulos et al., 2009b;Meynard et al., 2009). Exogenous BMP6 by intraperitoneal injection increases hepcidin expression and iron metabolism *in vivo* (Andriopoulos et al., 2009b). The expression of BMP6, the inducer of hepcidin expression, is sensitive to iron level in the serum. HJV, as a co-receptor to BMP6, however, does not regulate BMP6

expression. *Hjv-/-* mice have increased levels of BMP6 mRNA and reduced hepcidin expression that causes severe iron overload (Zhang et al., 2010). The increase of BMP6 mRNA level results from iron overload but not Hjv depletion, because increase in iron overload in wild type mice also results in a similar level of increase of BMP6 mRNA (Kautz et al., 2008a). This also indicates that BMP6 expression alone cannot induce hepcidin expression in the absence of Hjv, which is consistent with the notion that HJV functions as a co-receptor.

HJV post-translational processing and significance

HJV protein has three predicted N-glycosylation sites. The muscle-expressed and liver-expressed HJV proteins are differentially glycosylated, but the functional significance for the differential glycosylation is unknown (Fujikura et al., 2011). Under reducing conditions, cell membrane localized HJV proteins have three major species of 50, 35, and 20kD respectively. Under non-reducing conditions, HJV only runs as a single 50kD band, which corresponds to the full-length form. The 35kD and 20kD species are from an intracellular autoproteolytic cleavage event at amino acid 172 in the vWF type D domain between aspartic acid and proline in an mildly acidic environment, and these two fragments are joined together by disulfide bonds to form a two-chain form (Zhang et al., 2005;Kuninger et al., 2006a). The autocleavage site is conserved in RGMa and RGMb and similar two-chain forms have been observed for RGMa and RGMb (Niederkofler et al., 2004; Figure 1.6). The membrane-bound two-chain species of Hjv is the predominant isoform (Kuninger et al., 2006a;Silvestri et al., 2007). Several JH disease causing mutations appear to affect the autocleavage event. These mutations include those that lie in or near the cleavage site, D172E, F170C and W191C, as well as one predominant disease causing mutation,

G320V (Silvestri et al., 2007;Pagani et al., 2008). It has been found that these mutant forms of HJV are retained in the endoplasmic reticulum and have low signaling activities. However, other disease causing mutations, such as G99V and C119F, that have no effect on HJV autocleavage but have reduced signaling activities do reach the plasma membrane (Silvestri et al., 2007). These studies suggest that HJV cleavage is essential for their export to the plasma membrane.

HJV also undergoes active release by furin proprotein convertase. The soluble form of cleaved HJV (s-HJV) has a molecular mass of ~40kD (Lin et al., 2008;Silvestri et al., 2008;Lin et al., 2008) . The cleavage site has been mapped to a RNRR sequence that is not conserved in RGMa and RGMb (Silvestri et al., 2008;Lin et al., 2008; Figure 1.6). Furin is predominantly localized in the trans-Golgi network (TGN) compartment, with only a small portion of it being on the plasma membrane (Thomas, 2002). Consistent with this, it has been shown that HJV cleavage occurs in the TGN after it is being endocytosed from the plasma membrane (Maxson et al., 2009). Furthermore, the binding of neogenin to HJV on the plasma membrane triggers the retrograde trafficking and subsequent release of soluble HJV, s-HJV (Zhang et al., 2008). s-HJV is present not only in transfected cell lines and cell lines that endogenously express HJV, but also in human and rat serum (Kuninger et al., 2004;Lin et al., 2005a;Kuninger et al., 2006b;Silvestri et al., 2007;Zhang et al., 2007).

s-HJV inhibits BMP signaling by binding to BMP2, BMP4, and BMP6 to suppress hepcidin expression *in vitro* (Babitt et al., 2006;Lin et al., 2007;Kuns-Hashimoto et al., 2008b;Andriopoulos et al., 2009a;Nili et al., 2010). Injection of s-HJV decreases hepatic hepcidin mRNA level *in vivo* (Babitt et al., 2007). s-HJV is hypothesized to compete with hepatocyte membrane-bound HJV for BMP

ligands, acting as a negative regulator of the BMP-mediated hepcidin expression. Because hepcidin expression is positively regulated by body iron load, it is reasonable to hypothesize that body iron load may negatively regulate the amount of s-HJV. In fact, s-HJV release is negatively regulated by iron level *in vitro* (Lin et al., 2005b;Silvestri et al., 2007;Zhang et al., 2007). In acutely iron-deficient rats, decreased serum Transferrin and Ferritin saturation is associated with decreased hepcidin expression and increased serum s-HJV level (Zhang et al., 2007). Therefore, s-HJV is a negative regulator of hepcidin expression in response to body iron status.

A serine protease Matriptase-2, encoded by the gene *TMPRSS6*, is also involved in regulating body iron load (Silvestri et al., 2008). In mice and zebrafish, matriptase-2 mutations cause hepatic hepcidin overexpression and iron-deficient anemia (Folgueras et al., 2008;Silvestri et al., 2008;Ramsay et al., 2009). In humans, mutations in *TMPRSS6* cause iron-refractory iron deficiency anemia (IRIDA), a familial anemia disorder. The relationship between matriptase-2 and HJV has been investigated. matriptase-2 suppresses HJV-induced hepcidin expression. Matriptase-2 and HJV are co-expressed in the liver (Velasco et al., 2002). Moreover, they directly bind to each other in Hela cells (Silvestri et al., 2008). The binding results in the cleavage of cellular HJV into many fragments that are released to the media (Silvestri et al., 2008). However, the *in vivo* situation is more complex. Liver membrane HJV level is decreased rather than increased in the *Tmprss6-/-* mutant mice (Krijt et al., 2011). The expression of matriptase can inhibit hepcidin expression that is induced by the addition of HJV, although it is controversial if matriptase also decreases BMP-induced hepcidin expression (Du et al., 2008;Silvestri et al., 2008). *In vivo* studies by making double mutants between *Hjv-/-* and *Tmprss6-/-* or a catalytic mutation

Tmprss6^{mask} showed that the double mutants exhibit the *Hjv-/-* mutant phenotypes, suggesting that HJV acts downstream of matriptase-2 in the regulation of hepcidin expression (Finberg et al., 2010;Truksa et al., 2009). The regulation of matriptase-2 expression has been investigated recently. *Tmprss6* mRNA expression is stimulated by iron loading treatment or BMP6 protein injection and is blocked by injection of neutralizing antibody against BMP6 in mice (Meynard et al., 2011). Since iron induces hepcidin expression, the upregulation of matriptase-2, a negative regulator of hepcidin, by iron indicates that matriptase-2 likely represents a negative feedback regulating mechanism to balance hepcidin overexpression induced by iron to maintain systemic iron homeostasis.

1.5 Neogenin

1.5.1 Introduction

Neogenin is a type I transmembrane protein (Chan et al., 1996). It is homologous to DCC (Deleted in Colorectal Cancer). Neogenin and DCC share 50% identify in amino acid sequence. They have highly conserved extracellular domains, including 4 Immunoglobuin (Ig) domains and 6 Fibronectin type III (FNIII) domains (Figure 1.7). The intracellular region is less conserved between them with only three motifs being conserved: P1, P2, and P3. DCC and neogenin have orthologs in *Drosophila*, zebrafish, *Xenopus*, chickens and mammals (Cole et al., 2007;Wilson and Key, 2007;Yamashita et al., 2007). In *C. elegans*, a single protein UNC-40 resembles both DCC and neogenin (Chan et al., 1996). In vertebrates, neogenin has a broad expression pattern compared to DCC. Neogenin is found in dividing neurogenic and gliogenic progenitors throughout

the embryonic and adult CNS (Vielmetter et al., 1994;Keeling et al., 1997;Fitzgerald et al., 2006a;Fitzgerald et al., 2006b;Fitzgerald et al., 2007). Outside of the nervous system, neogenin protein is found in many developing and adult tissues, including heart, gut, lung, liver (Gad et al., 1997;Rodriguez et al., 2007b), and limb buds (Hong et al., 2012). In contrast, the expression of DCC is primarily restricted to the developing nervous system.

1.5.2 Neogenin and DCC as receptors for netrins

Both neogenin and DCC are receptors for netrins, which are a class of axon guidance molecules (Figure 1.7). The netrin-binding site on DCC has been mapped to the fourth and fifth FNIII domains (FNIII-4, FNIII-5) (Geisbrecht et al., 2003;Kruger et al., 2004). Because of the high degree of similarity between neogenin and DCC in their extracellular region, neogenin is expected to bind to netrin via similar domains. In fact, the binding affinity between netrin and neogenin was shown to be around 2nM (Wang et al., 1999b), which is simiar to the DCC-netrin binding affinity (Keino-Masu et al., 1996). DCC-netrin interactions are required for chemoattractant and chemorepellent responses (Kennedy et al., 1994;Colavita and Culotti, 1998;Hong et al., 1999;Wang et al., 1999a;Qin et al., 2007). Although less well studied, neogenin has a netrin-dependent role in mediating chemoattractant responses to netrin in supraoptic axons in *Xenopus* forebrain (Wilson and Key, 2006). In addition to axon guidance, both DCC and neogenin regulate cell-cell adhesion and tissue organization through interactions with netrins (Srinivasan et al., 2003;Kang et al., 2004;Park et al., 2004;Jarjour et al., 2008;Lejmi et al., 2008). The signal transduction cascades initiated upon DCC-netrin interaction have been mapped out in some details, which include intracellular tyrosine kinase, second messenger, and rho GTPase pathways

(Gomez and Zheng, 2006;Lai Wing Sun et al., 2011;Lin and Holt, 2007;Moore et al., 2007;Rajasekharan and Kennedy, 2009;Round and Stein, 2007), intracellular signaling downstream of neogenin-netrin is poorly understood. Limited evidence suggests that neogenin turns on similar set of downstream signaling events to that of DCC (Li et al., 2004;Liu et al., 2004;Xie et al., 2005;Ren et al., 2008). However, it is noted that in at least several cases neogenin and DCC have differential downstream effects and interaction partners. For example, DCC and neogenin, as cargos for myosin X, promote different movements of myosin X (Liu et al., 2012); different to DCC, neogenin fails to interact to PLCγ but binds to Src homology-2 containing inositol-5-phosphatase 1 (SHIP1) in the cytoplasmic region in second messenger pathways (Xie et al., 2006;Ren et al., 2008).

1.5.3 Neogenin is a receptor for RGMs

Neogenin is not only a receptor for netrins, but also a receptor for RGMs (Figure 1.7). Neogenin, but not DCC, directly binds to all three RGM family members *in vitro*. The RGMa-binding site is known to reside within the FNIII domains of neogenin (Rajagopalan et al. 2004). Further mapping of the binding region indicates that neogenin's FNIII-(3-4) region is critical for RGMa binding (Tassew et al., 2012). Interestingly, the binding affinity of RGMa to neogenin is about 10-fold higher than that of netrin to neogenin, suggesting that RGMa may be the primary ligand for neogenin in the CNS (Rajagopalan et al. 2004). Because the binding of RGMa to neogenin is unaffected by netrin-1 binding, these two ligands probably have distinct binding sites (Rajagopalan et al. 2004). Unlike for RGMa, the binding site for HJV is mapped to the FNIII-5, FNIII-6 and the proximal juxtamembrane region of neogenin, in which FNIII-6 mediates most of the interaction, while FNIII-5 alone does not bind HJV (Yang et al.,

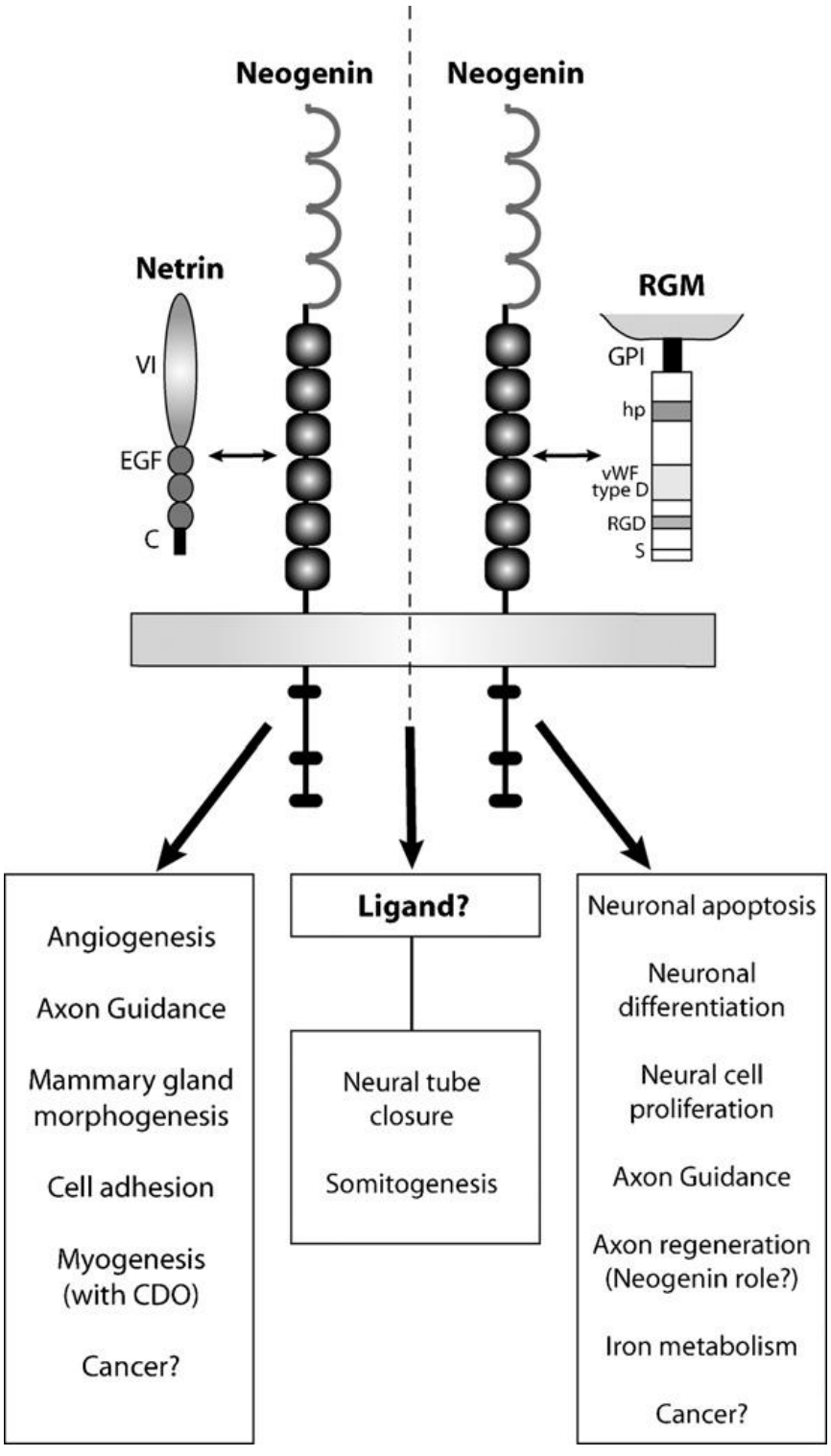
2008a;Yang et al., 2011). The crystal structure of FNIII-5 and FNIII-6 of neogenin has been solved (Yang et al., 2011). By comparing neogenin FNIII-6 domain (that binds to HJV) to neogenin FNIII-5 domain and DCC FNIII-6 domain (that do not bind to HJV), two loops in neogenin FNIII-6 domain are predicted to mediate HJV-neogenin interaction (Yang et al., 2011). Efforts have been made to map the neogenin binding site in RGMa. In one study, neogenin binding sites in RGMa were shown to contain amino acids (aa) 259-295 (Itokazu et al., 2012). However, another study indicates a surprising result that the two non-overlapping RGMa N- and C- fragments, which are resulted from RGMa autocleavage, could both bind to neogenin fragments. Although the neogenin binding sites in HJV haven't been mapped, the most prevalent human JH mutation G320V abolishes neogenin-HJV interaction. It may be due to the abolishment of direct interaction or the induction of a conformational change in HJV that affects neogenin-HJV interaction (Kuns-Hashimoto et al., 2008a). It is important to note that HJV can simutaneously bind neogenin and BMP2, suggesting that the binding sites for each protein in HJV are not overlapping (Yang et al., 2008a).

Neognenin-RGMa interaction was first identified to be required for the chemorepulsion of temporal retinal axons in the anterior tectum of the chick (Monnier et al., 2002). Later, neogenin-RGMa interaction was recognized in many of the processes in cultured cells and mice that RGMa regulates, including axon tract formation in the embryonic brain (Matsunaga et al., 2004;Monnier et al., 2002;Tassew et al., 2008), neural tube closure (Kee et al., 2008), neuronal differentiation (Matsunaga et al., 2006), cell survival (Koeberle et al., 2010), cell migration, and cell-cell adhesion (Lah and Key, 2012; Figure 1.7). Besides RGMa, neogenin also directly binds to RGMb and HJV. RGMb-neogenin interaction controls aggregation and migration of a population of dentate precursor cells *in*

vitro and in mice (Conrad et al., 2010). The downstream signaling of RGMa-neogenin binding appears to be different than that of netrin-neogenin binding. It has been shown that RGMa-neogenin-dependent growth cone collapse is mediated by rhoA activation of rhoA kinase (ROCK; Conrad et al., 2007;Hata et al., 2006). The activated rhoA kinase can phosphorylate the light chain of myosin II in an RGMa-dependent manner, leading to F-actin reduction and growth cone collapse and subsequent neurite retraction/outgrowth inhibition (Kubo et al., 2008).

Figure 1.7: **The roles of neogenin in mediating netrin and RGM signaling**

Neogenin is a transmembrane protein with 4 Ig-like domains (half circles), 6 FNIII domains (black boxes) and three conserved intracellular motifs: P1, P2, and P3 (black bars). Neogenin is a multifunctional receptor that mediates signaling from the soluble protein netrin and the GPI-anchor proteins RGMs for different cellular functions. Modified from Wilson and Key, 2007.


Neogenin
Neogenin
Netrin
VI
EGF
C
RGM
GPI
hp
vWF
type D
RGD
S
Angiogenesis
Axon Guidance
Mammary gland morphogenesis
Cell adhesion
Myogenesis (with CDO)
Cancer?
Ligand?
Neural tube closure
Somitogenesis
Neuronal apoptosis
Neuronal differentiation
Neural cell proliferation
Axon Guidance
Axon regeneration (Neogenin role?)
Iron metabolism
Cancer?

1.5.4 Neogenin functions in the BMP signaling pathway

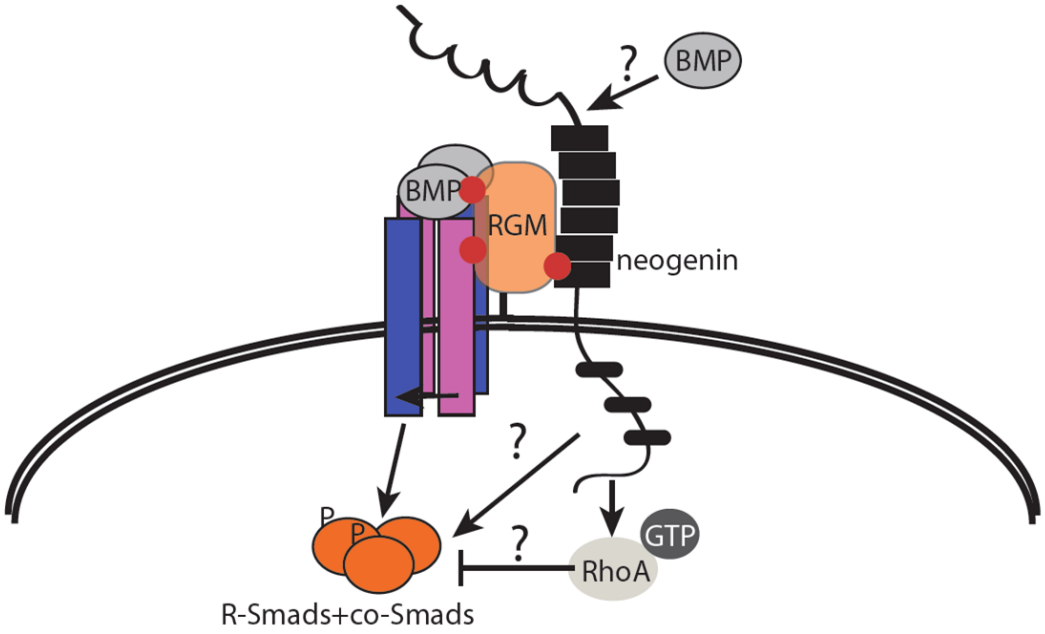
The broad presence of neogenin outside of the nervous system suggests that it may have other functions. One of its functions is modulating BMP signaling. Neogenin-deficeint mice that lack functional neogenin have many defects and could survive for only one month (Bae et al., 2009;Lee et al., 2010). Neogenin-deficient mice have reduced hepatic hepcidin expression, severe iron overload, and reduced BMP signaling, which are similar phenotypes as those seen in *Hjv-/-* mice. Reduction of hepcidin expression has been observed in hepatoma cells by knocking down neogenin with or without ectopic HJV expression (Zhang et al., 2009). Disruption of neogenin-HJV interaction by a soluble neogenin fragment suppresses BMP4 induced hepcidin expression, suggesting that neogenin-HJV interaction regulates BMP4-HJV-induced hepcidin expression (Zhang et al., 2009). It is important to note neogenin and HJV share similar expression patterns in the liver, suggesting they may function in the same cells *in cis* (Lee et al., 2010). Neogenin-deficient mice are also impaired in endochondral ossification. It has been shown that neogenin regulates endochondral bone formation through canonical BMP signaling in a cell autonomous fashion (Zhou et al., 2010).

The mechanism of how neogenin modulates BMP signaling is controversial and not well understood. It may also be dependent on cellular context. First, neogenin-HJV interaction has been shown to regulate HJV secretion *in vitro*. Knockdown of endogenous neogenin in differentiated myoblast and hepatoma cells abolished HJV secretion (Zhang et al., 2005;Zhang et al., 2007). Inhibition of neogenin-HJV interaction by adding soluble neogenin to the media also reduced HJV secretion (Zhang et al., 2008). The direct effects on BMP signaling and hepcidin expression haven't been determined in the above

studies; however, we know that soluble HJV is an antagonist of BMP signaling and hepcidin expression. As a prediction, enhancing the secretion of a BMP signaling antagonist by neogenin would result in decrease in the BMP signaling. There is one study showing the opposite result in which neogenin stabilizes HJV and inhibits HJV secretion in HEK293 cells (Lee et al., 2010). This latter piece of evidence is consistent with the above *in vivo* observations that neogenin positively modulates hepcidin expression and BMP signaling. Second, neogenin has been shown to promote the entry of BMP receptors to the lipid rafts microdomains, which is necessary for BMP signaling. Because neogenin on its own has little direct interaction with the BMP receptors, RGMa appears to act as a bridge protein for neogenin to associate with BMP receptors (Zhou et al., 2010). Third, in C2C12 cells, neogenin was found to negatively regulate the BMP2-induced processes of osteoblastic differentiation. It was found that by directly binding to BMP ligands and activating RhoA independent of BMP receptors, neogenin antagonizes phosphorylation of Smads1/5/8 (Hagihara et al., 2011). The negative role of neogenin in BMP signaling may be context dependent and requires support from *in vivo* analysis. Based on these different pieces of published data, I illustrate on Figure 1.8 a model of how neogenin might function in modulating BMP signaling.

Figure 1.8: A model for how neogenin might function in BMP signaling

Neogenin forms a complex by directly binding to the RGM protein, which directly binds to BMP receptors at the extracellular region. The formation of the complex can positively modulate BMP signaling, possibly through recruiting BMP ligands to the lipid raft microdomains. It is not clear if the intracellular domain of neogenin plays a role here. Neogenin also directly binds to BMP ligands via its ectodomain. BMP binding activates RhoA which inhibits phosphorylation of Smads, therefore, repressing BMP signaling. The mechanism and *in vivo* relevance of this repression is not clear. Red dots represent direct physical contacts.


?
BMP
BMP
RGM
neogenin
?
?
GTP
RhoA
P
P
R-Smads+co-Smads

1.5.5 Summary

Neogenin is a multifunctional receptor with a broad expression pattern. It mediates netrin and RGM signaling in axon guidance, neuronal development, apoptosis, and survival through differential downstream signaling. Neogenin has also been shown to positively modulate BMP signaling in regulating hepcidin expression to balance iron metabolism and to affect endochondral ossification. In cell cultures, however, neogenin appears to have both positive and negative roles in BMP signaling. The direct interaction between neogenin and RGM can affect the stability of RGM protein on the plasma membrane and recruit BMP receptors to lipid rafts. Neogenin can also directly interact with the BMP ligands, instead of promoting BMP signaling, suppressing BMP signaling by activating RhoA. There are many unanswered questions: Is modulation of BMP signaling involved in the neuronal roles of neogenin? How is BMP signaling mediated by neogenin *in vivo*: does neogenin function by affecting the stability of RGM proteins; does neogenin always partner with RGM in modulation of BMP signaling; what is the *in vivo* relevance of neogenin as a negative regulator of BMP signaling; does neogenin modulate BMP signaling by acting through any intracellular transducers? Furthermore, as a widely expressed protein, does neogenin have other roles? Neogenin has recently been shown to be involved in digit patterning by regulating the Sonic Hedgehog pathway (Hong et al,. 2012). The mechanisms are currently unknown. Based on the many unanswered questions and the somehow conflicting observations in cell culture studies, it is obvious that *in vivo* mechanistic studies are in demand and we may be able to shed light on its function via studies in model organisms such as *C. elegans*.

1.6 Dissertation Outline

This dissertation has two major themes, including the description of a genetic network important for M lineage development that centered around a sole *C. elegans* SoxC protein, and the characterization of RGM and neogenin proteins as modulators of the Sma/Mab BMP-like pathway.

Chapter 2 describes the identification and characterization of *sem-2/SoxC* as a pro-proliferation factor that is both required and sufficient for the SM (Sex Myoblast) fate. My work also revealed the regulatory network that spatiotemporally controls *sem-2* expression.

Chapter 3 describes the characterization of the *C. elegans* RGM protein DRAG-1 as a tissue specific BMP coreceptor.

Chapter 4 describes recent advances the power of DRAG-1 and *the C. elegans* to study vertebrate RGM protein function.

Chapter 5 describes the characterization of neogenin/UNC-40 as a positive modulator of the Sma/Mab signaling pathway.

Chapter 6 summarizes this work and discusses my thoughts on future directions following my studies.

Appendix I describes the relationship between *sem-2/SoxC* and *sem-4/SALL* in specifying the SM cells.

Chapter 2: The *C. elegans* SoxC protein SEM-2 opposes differentiation factors to promote a proliferative blast cell fate in the postembryonic mesoderm[1]

2.1 INTRODUCTION

The group C Sox proteins are Sry-related HMG box (Sox)-containing transcription factors (Wegner, 2010). Vertebrates contain three highly conserved SoxC genes, Sox4, Sox11 and Sox12, which play important roles in development, including cell differentiation, proliferation and survival (Penzo-Mendez, 2009). Increasing evidence has also shown that many tumor types in humans are associated with significantly elevated level of SoxC gene expression, suggesting that mis-regulation of the SoxC genes may contribute to tumor formation (Penzo-Mendez, 2009). However, the molecular mechanisms underlying these multiple functions of the SoxC genes are not fully understood, nor are the molecular events that regulate the expression of SoxC genes.

The nematode *C. elegans* contains a single SoxC gene, C32E12.5 (Phochanukul and Russell, 2010), providing an opportunity to determine the functions of SoxC genes at single cell resolution. In this study, we show that C32E12.5 is the *sem-2* gene that is widely expressed and essential for *C. elegans* embryonic development. Furthermore, *sem-2* plays a critical role in a binary fate decision in the postembryonic mesoderm, the M lineage. The M

[1] This chapter was published in *Development* (Tian C, Colledge C, Shi H, Stern MJ, Waterston RH, Liu J. Development. 2011 Mar;138(6):1033-43.) and is reprinted with permission. Mapping was performed by Jun Liu, C. Colledge, who was student of R.H. Waterston, molecular cloning was performed by H. Shi, mutation isolation and initial characterization was performed by M.J. Stern.

mesoblast is an embryonically-born pluripotent precursor cell (Fig. 1A). During hermaphrodite larval development, the M cell first divides dorsoventrally, then left-right, and then twice anterioposteriorly to produce 16 cells (Sulston and Horvitz, 1977). Two of them, M.vlpa and M.vrpa, divide one more time. The anterior cells from these final divisions become sex myoblasts (SMs), the precursors to all the non-striated egg-laying vulval and uterine muscles. The two posterior cells differentiate into striated body wall muscles (BWMs). The SMs then migrate to the future vulval region and further proliferate to produce 8 vulval muscles (vms) and 8 uterine muscles (ums). We show here that *sem-2* is required for the binary fate decision between the SMs and BWMs, and is both necessary and sufficient to promote the proliferative SM fate as opposed to the differentiated BWM fate. This specific function and expression of *sem-2* in the M lineage is under the direct regulation of Hox/PBC proteins, MAB-5, LIN-39 and CEH-20. This finding is intriguing in light of the oncogenic roles of Hox and PBC factors (Shah and Sukumar, 2010), suggesting the possibility of Hox regulation of SoxC genes during tumorigenesis.

2.2 MATERIAL AND METHODS

2.2.1 C. elegans strains

The following strains were used in this study: LG I: *sem-2(n1343)*; *sem-2(ok2422)/hT2[qIs48]*; *sys-1(q544)/hT2[qIs48]*. LG II: *hlh-1(cc561ts)*. LG III: *fozi-1(cc609)*; *mab-5(e1239) lin-39(n1760)/hT2[qIs48]*; *ceh-20(os39)/hT2[qIs48]*; *unc-32(e189) lin-12(n676n930ts)*. LG X: *sma-9(cc604)*. Analyses were performed at 20°C, unless otherwise noted.

Integrated lines:

LW0081: *ccIs4438 (intrinsic CC::gfp)* III; *ayIs2 (egl-15::gfp)* IV; *ayIs6 (hlh-8::gfp)* X

LW1066: *jjIs1066[pJKL705.1(hlh-8p::mRFP+unc-119(+)]?; unc-119(ed4)* III

LW1639: *mab-5(e1239) lin-39(n1760)* III*/hT2[qIs48]* (I;III); *ayIs6(hlh-8p::gfp)* X

LW2466: *jjIs1475(myo-3::rfp)* I; *ccIs4438 (intrinsic CC::gfp)* III; *ayIs2 (egl-15::gfp)* IV; *ayIs6 (hlh-8::gfp)* X

2.2.2 Isolation, genetic and molecular analysis of sem-2 alleles

sem-2(n1343) was identified in a Tc1 mutagenesis screen (Desai et al., 1988). It was mapped using three factor crosses into a small region between *unc-63* and *spe-11*, 0.2 map units to the right of *unc-63*. Cosmids spanning that region were tested for rescuing activity. Two cosmids, F47D3 and C32E12, both rescued the egg laying defect of *n1343*. By further truncating the overlapping region of the cosmids, the rescuing activity was mapped to C32E12.5. To identify the molecular lesion of *n1343*, PCR reactions were performed using primers designed to span the genomic region of C32E12.5. Primer pairs JKL-620 and JKL-621 amplified a fragment bigger than expected. Further PCR and sequencing identified a Tc1 insertion in the first intron of C32E12.5 (between -9186 and -9185, Figure 2A).

2.2.3 Plasmid constructs and transgenic lines

Fosmid WRM0623cE02 (Geneservice Ltd) was used to generate the *gfp::sem-2* translational fusion construct via recombineering (Warming et al., 2005). *gfp* sequences from the Fire lab vector pPD95.75 were inserted immediately after the ATG start codon. Recombineering was also used to insert Tc1 into the *gfp::sem-2*-containing fosmid generated above between -9186 and -9185. The insertions were verified by PCR and sequencing.

sem-2 cDNA from *yk657g12* was used to generate pJKL776 (*hlh-8p::sem-2 cDNA::unc-54 3'UTR*).

A 4554bp fragment of the *sem-2* 1st intron (-11650 to -7097) was cloned into L3135 of the Fire lab vector kit to generate pCXT12, which was subsequently used to generate deletion constructs pCXT18-22 and pCXT26-33. pCXT33 was used to generate pCXT97-99 and pCXT173, carrying mutations in the PBC/Hox binding site. *sys-1*(T23D8.9) and *let-381*(F26B1.7) RNAi constructs were obtained from the Ahringer RNAi library provided by Geneservice Ltd (Kamath et al., 2003). *sem-2(RNAi)* construct pCXT9 was made by subcloning the *sem-2* cDNA fragment from *yk657g12* into L4440 (Timmons and Fire, 1998). pNMA49 and pNMA50 were used for knocking down *fozi-1* and *mab-5*, respectively (Amin et al., 2009).

Transgenic lines were generated using the plasmids pRF4 (Mello et al., 1991), pJKL449 (*myo-2p::gfp::unc-54 3' UTR*) (Jiang et al., 2009) or LiuFD61 (*mec-7p::mRFP)* (Amin et al., 2009) as markers.

2.2.4 RNAi

The T7 Ribomax RNA Production System (Promega) was used to generate *sem-2* dsRNA using *yk404e6* as a template. Synchronized L1 animals expressing various M lineage GFP markers were soaked in the dsRNA solution at 20˚C for 24-48 hours following the protocol of Maeda and colleagues (Maeda et al., 2001). Animals were allowed to recover at 20˚C and adult worms were scored for M lineage phenotypes. Water was used as a soaking control. For *fozi-1(RNAi)* (Amin et al., 2007)*, let-381(RNAi)* (Amin et al., 2010)*, mab-5(RNAi)* and *sys-1(RNAi)* (Amin et al., 2009)*,* synchronized L1 animals expressing different M lineage markers were plated on HT115(DE3) bacteria expressing dsRNA for the gene of interest. Bacteria for ingestion were prepared as described by Kamath and Ahringer (2003). RNAi by ingestion was performed at 25˚C and animals were scored for M lineage phenotypes 24-48h after plating.

2.2.5 Immunofluorescence staining

Animal fixation, immunostaining, microscopy and image analysis were performed as described previously (Amin et al., 2007). Guinea pig anti-FOZI-1 (Amin et al., 2007) (1:200) and goat anti-GFP (Rockland Immunochemicals; 1:1000) were used.

2.2.6 Electrophoretic Mobility Shift Assays (EMSA)

6xHis-tagged LIN-39 and CEH-20 fusion proteins were purified as previously described (Liu and Fire, 2000). Complementary single-stranded DNA

oligos were 3'-end labeled with biotin using the Biotin 3' End DNA Labeling Kit (Pierce) and annealed at room temperature for one hour. Gel shift reactions and detection were performed using the LightShift Chemiluminescent EMSA Kit (Pierce). Oligonucleotides used are: wt: CXT216/217, canonical: CXT218/219, mut 1: CXT220/221, mut 2: CXT222/223, mut 3: CXT224/225.

2.3 RESULTS

2.3.1 sem-2(n1343) mutants exhibit a fate transformation of the proliferating SMs to differentiated BWMs

sem-2(n1343) was identified in a Tc1 insertion mutagenesis screen for egg-laying defective (Egl) mutants in the *mut-2(r459)* background (Desai et al., 1988). Hermaphrodites homozygous for *n1343* are 100% Egl (n>200,), lack all vulval and uterine muscles required for egg-laying, as monitored by DIC and polarized light microscopy, and fail to express egg-laying muscle specific reporters, such as *egl-15::gfp* and *ceh-24::gfp* (expressed in type I vulval muscles, VM1s (Harfe and Fire, 1998b), or in all the egg-laying muscles (Harfe and Fire, 1998)) (Figure 2.1C-F, data not shown). To determine the basis for the missing egg-laying muscles, we followed the M lineage in *n1343* hermaphrodites using both Normaski microscopy and the *hlh-8::gfp* reporter that marks all undifferentiated cells of the M lineage (Harfe et al., 1998b). In *n1343* hermaphrodites, the SM mother cells, M.v(l/r)pa as in wild type give rise to two daughter cells, M.v(l/r)paa and M.v(l/r)pap. However, the anterior daughters,

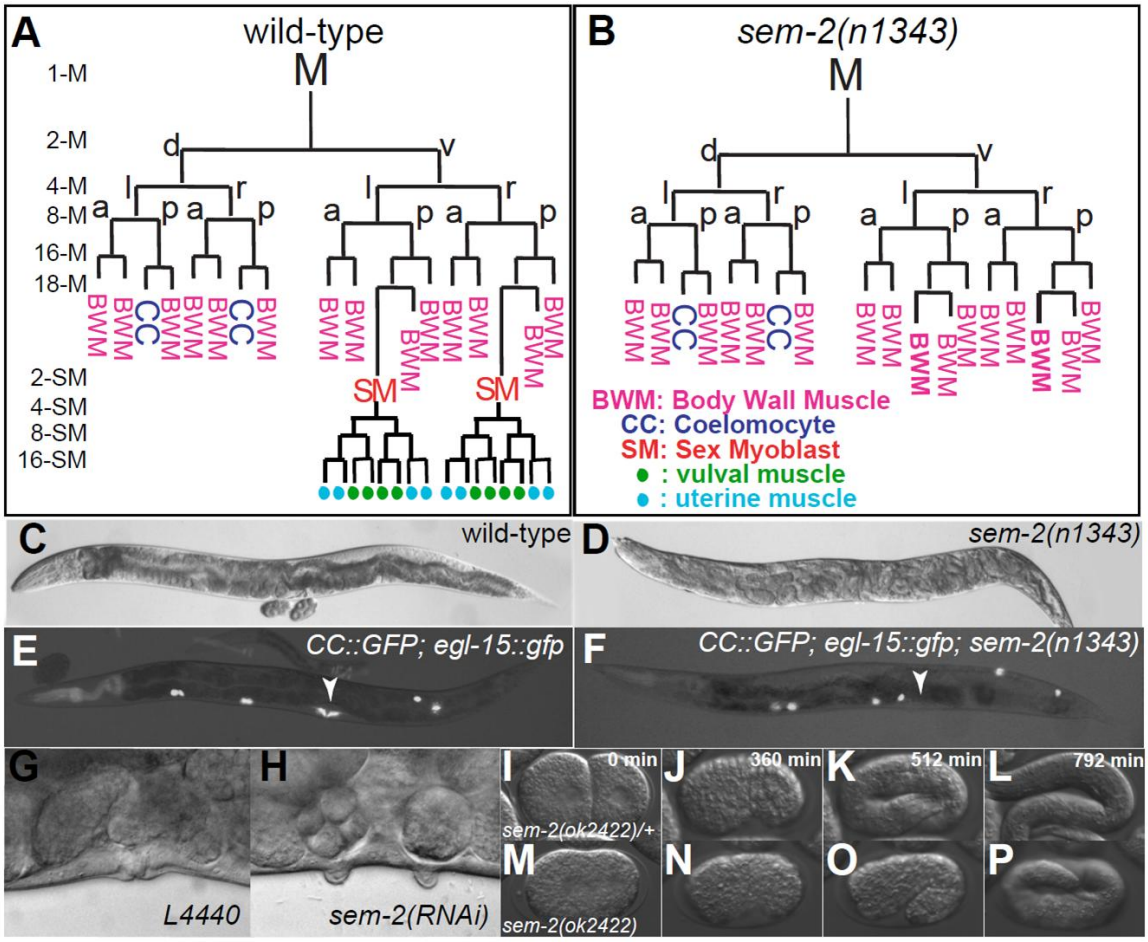
normally the SMs, exhibit the morphology and identity of their sister cells, the body wall muscles (BWM, Figure 2.1A,B and data not shown). These cells also fail to migrate anteriorly, never divide, and never differentiate into the egg-laying muscles. Therefore, *n1343* hermaphrodites exhibit a SM to BWM fate transformation. The M lineage phenotype of *n1343* animals may reflect a complete removal of *sem-2* function in the M lineage because *n1343* over a deficiency of the *sem-2* region, *sDf4*, yielded the same M lineage phenotype as homozygous *n1343* animals (data not shown). The effect of the *sem-2(n1343)* mutation on SM development appears to be sex specific, since the SMs appear unaffected in *n1343* males, and *n1343* males mate efficiently (data not shown).

2.3.2 sem-2 encodes the sole C. elegans group C HMG/SRY box-containing protein

We mapped *sem-2(n1343)* to C32E12.5 based on 3-factor mapping and cosmid rescue (see Materials and Methods). PCR and sequencing analyses showed that *n1343* animals contain a Tc1 transposon insertion located 9185bp upstream of the predicted ATG of C32E12.5 (Figure 2.2A). RNAi of C32E12.5 in wild-type worms by soaking L1s caused 98% (n=500) of worms to be Egl and missing *egl-15::gfp* expression. RNAi treated worms in the same experiment also exhibited a multivulvae phenotype, with 48% being bi-vulvae and 5% having three vulvae (n=81, Figure 2.1G,H and data not shown).

Figure 2.1: *sem-2* is required for proper embryonic and postembryonic development.

(A) The wild-type postembryonic M lineage showing all differentiated cell types that arise from M [modified from Sulston and Horvitz (1977)]. (B) The M lineage of *sem-2(n1343)* mutants. (C, D) DIC images of a wild-type (C) and a *sem-2(n1343)* (D) adult hermaphrodites. All images presented hereafter are oriented with dorsal side up and anterior to the left, unless noted otherwise. (E, F) A wild-type (E) and a *sem-2(n1343)* animal (F) with *CC::gfp* and *egl-15::gfp*. Arrow heads indicate the vulva. Both wild-type and *n1343* animals have four embryonically-derived and two M-derived CCs. (G, H) Vulval phenotypes of *sem-2(RNAi)* animals (H) as compared to *L4440* control RNAi animals (G). (I-P) frames from time-lapse movies of *sem-2(ok2422)/+* (I-L) and *sem-2(ok2422)* (M-P) worms. Time 0 is when the movie was started. Four time points were selected to represent different embryonic stages. M: M mesoblast, d: dorsal, v: ventral, a: anterior, p: posterior.


A
wild-type
1-M
2-M
4-M
8-M
16-M
18-M
2-SM
4-SM
8-SM
16-SM
M
d
v
l
r
a
p
BWM
CC
SM
B
sem-2(n1343)
BWM: Body Wall Muscle
CC: Coelomocyte
SM: Sex Myoblast
: vulval muscle
: uterine muscle
C
wild-type
D
sem-2(n1343)
E
CC::GFP; egl-15::gfp
F
CC::GFP; egl-15::gfp; sem-2(n1343)
G
L4440
H
sem-2(RNAi)
I
0 min
J
360 min
K
512 min
L
792 min
sem-2(ok2422)/+
M
sem-2(ok2422)
N
O
P

The *n1343* mutation does not appear to be a null allele of *sem-2,* since a deletion allele of *sem-2, ok2422,* generated by the *C. elegans* knockout consortium, resulted in 100% embryonic lethality (N>100). *ok2422* has most of the C32E12.5 coding sequences deleted and is likely a null allele (Figure 2.2A). Time-lapse video microscopy showed that *ok2422* homozygous embryos developed normally until late-comma stage (n=5, Figure 2.1I-P). However, these embryos were severely delayed in the elongation process, resulting in noticeable morphological defects such as enlarged heads relative to other 1½ fold embryos (Figure 2.1O). They eventually all reached to, and were arrested at, the 3-fold stage (Figure 2.1L,P). A fosmid (WRM0623cE02) containing C32E12.5 rescued the embryonic lethality of *ok2422.* RNAi soaking of L4s also resulted in 100% embryonic lethality (N>200). Thus, C32E12.5 is an essential gene required in embryogenesis, vulval and M lineage development, and *n1343* is a partial loss-of-function allele of C32E12.5 specifically affecting M lineage development.

Sequencing of the cDNA clones for C32E12.5 showed that the C32E12.5 locus is alternatively spliced with two different splice isoforms. Splice form 1 is represented by the cDNA clone *yk657g12*, in which two small exons separated by two large introns, 4.5kb (containing the Tc1 insertion in *n1343* animals) and 7kb respectively, are located upstream of the ATG-containing exon (Figure 2.2A). Splice form 2 is represented by two independent cDNA clones *yk1577b07* and *yk1661e08*, which have a SL1 splicing leader sequence trans-spliced to the ATG-containing exon (Figure 2.2A). Both splice isoforms are predicted to encode the same open reading frame that contains 404 amino acids. The predicted

SEM-2 protein contains a DNA-binding domain, the SRY/HMG box (residue 92-156), and a C terminal serine rich region (residue 334-354) that is predicted to be the transcriptional activation domain (Figure 2.2B). Based on the homology in the SRY/HMG region, SEM-2 is most similar to group C Sox proteins including Sox4, Sox11 and Sox12 in vertebrates and Sox14 in Drosophila (Bowles et al., 2000).

2.3.3 SEM-2/SoxC is a nuclear protein expressed in the SM precursors and their descendants

To understand how *sem-2* functions, we generated a N-terminal *gfp::sem-2* translational fusion and examined its expression pattern. This *gfp::sem-2* translational fusion is functional because it rescued the Egl phenotype of *n1343* (1/1 line) and the embryonic lethality of *ok2422* (2/2 lines). The GFP::SEM-2 protein was nuclear localized, consistent with its predicted role as a transcription factor (Figure 2.2C,E,F,I,L,O).

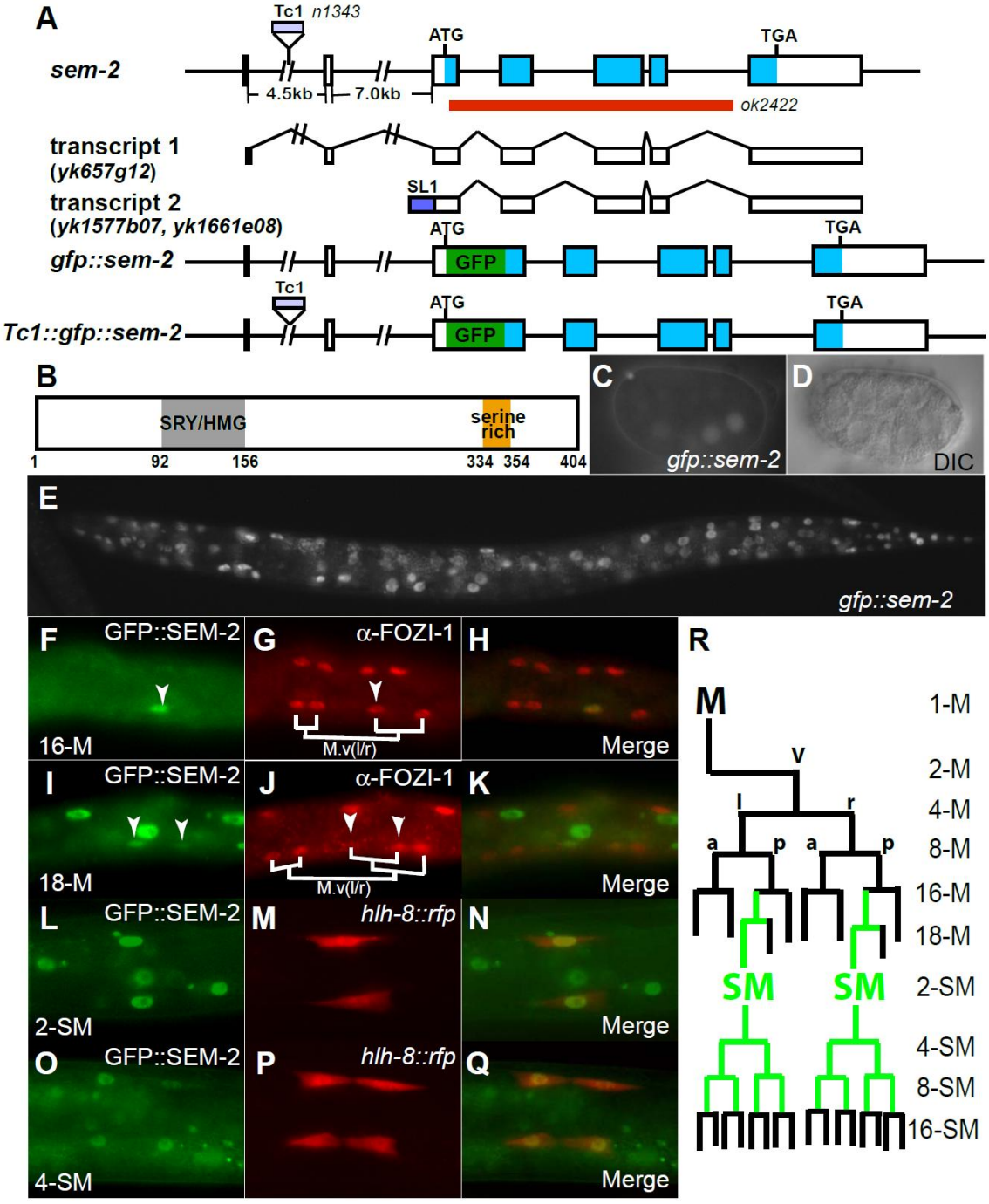
Expression of *gfp::sem-2* was first detectable in a subset of cells of the E and MS lineages in early gastrulating-stage embryos (Figure 2.2C, D). A similar expression pattern has been reported by Broitman-Maduro et al. (2005) using a transcriptional reporter of *sem-2*. The *gfp::sem-2* expression persisted through embryonic and larval development in many cell types, including vulval, hypodermal and intestinal cells (Figures 2.2E,2.3D).

To determine the expression pattern of *gfp::sem-2* in the M lineage, we also labeled the M lineage cells with *hlh-8p::rfp* or anti-FOZI-1 immunostaining (Amin et al., 2007). *gfp::sem-2* expression in the M lineage was first detectable at

the 16-M stage in the SM mother cells, M.v(l/r)pa (Figure 2.2F-H), and remained in both of their daughter cells, M.v(l/r)paa and M.v(l/r)pap (Figure 2.2F-H). The presence of GFP::SEM-2 in M.v(l/r)pap was transient: GFP::SEM-2 was not detectable after M.v(l/r)pap differentiated into BWMs. However, GFP::SEM-2 persisted in the nuclei of the SM cells and all their descendants until the 8-SM stage, and became undetectable at the 16-SM stage (Figure 2.2L-Q). The expression pattern of GFP::SEM-2 is summarized in Figure 2.2R. Thus, *sem-2* expression is turned on in the SM mother cells and is retained in the SMs and their descendants.

Figure 2.2: *sem-2* gene structure and expression pattern.

(A) *sem-2* gene structure (not drawn to scale) showing the molecular lesions of the *sem-2(n1343)* and *sem-2(ok2422)* mutations, two splice forms identified by cDNAs, and two translational reporters: *gfp::sem-2* and *Tc1::gfp::sem-2*. Exons are in boxes. The coding region is in blue. (B) A schematic of the predicted SEM-2 protein. SEM-2 contains a SRY/HMG box and a serine rich region. (C, D) *gfp::sem-2* expression (C and the corresponding DIC image (D)) is first detectable in embryos at the beginning of gastrulation. (E) Expression of *gfp::sem-2* in a transgenic larva. (F-Q) Expression of *gfp::sem-2* in the M lineage. (F-K) Double labeling of wild-type animals with anti-GFP antibodies (F, I) and anti-FOZI-1 antibodies (G, J) and the corresponding merged images (H, J) at the 16-M (F-H) and 18-M (I-K) stages, showing the expression of *gfp::sem-2* in the SM mother cell (M.vlpa) and the SM cell (M.vlpaa). Expression is also seen in the equivalent cells on the right side (not shown). (L-Q) Ventral views of wild-type animals carrying *gfp::sem-2* (L, O) and *hlh-8::rfp* (M, P) and the corresponding merged images (N, Q) at the 2-SM (L-N) and 4-SM (O-Q) stages. (R) Summary of GFP::SEM-2 expression in the M lineage. GFP::SEM-2-expressing cells are in green.


A
Tc1 n1343
ATG
TGA
sem-2
4.5kb
7.0kb
ok2422
transcript 1
(yk657g12)
transcript 2
(yk1577b07, yk1661e08)
SL1
gfp::sem-2
GFP
Tc1::gfp::sem-2
B
SRY/HMG
serine rich
1
92
156
334
354
404
C
gfp::sem-2
D
DIC
E
gfp::sem-2
F
GFP::SEM-2
16-M
G
α-FOZI-1
M.v(l/r)
H
Merge
I
GFP::SEM-2
18-M
J
α-FOZI-1
M.v(l/r)
K
Merge
L
GFP::SEM-2
2-SM
M
hlh-8::rfp
N
Merge
O
GFP::SEM-2
4-SM
P
hlh-8::rfp
Q
Merge
R
M
1-M
v
2-M
l
r
4-M
a
p
a
p
8-M
16-M
18-M
SM
SM
2-SM
4-SM
8-SM
16-SM

2.3.4 M lineage expression of sem-2 is specifically disrupted in n1343 mutants

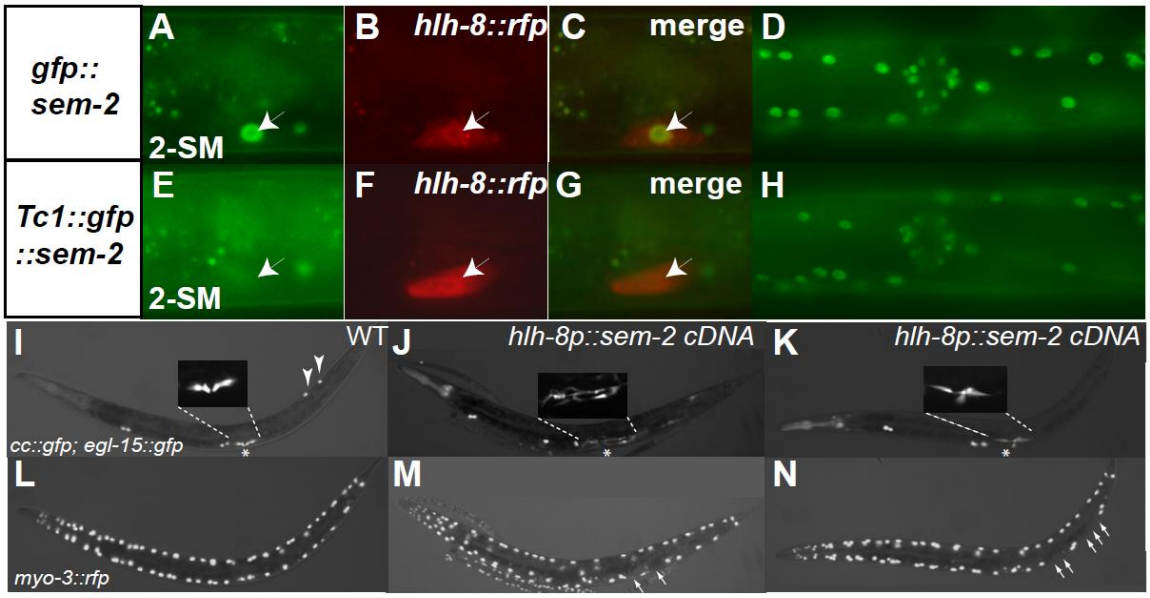
Since *sem-2* is an essential gene, the exclusive M lineage defects observed in *sem-2(n1343)* mutants suggest that the Tc1 transposon insertion may specifically affect *sem-2* expression in the M lineage. To test this hypothesis, we introduced the Tc1 transposon back into the functional *gfp::sem-2* translational fusion, at the same insertion site as found in *n1343* mutants (Figure 2.2A), and examined the function and expression of this reporter: *Tc1::gfp::sem-2. Tc1::gfp::sem-2* rescued the embryonic lethality of *ok2422* mutants (1/1 line). However, it failed to rescue the Egl phenotype of *n1343* mutants (4/4 lines, n>100). Thus *Tc1::gfp::sem-2* is functional outside of the M lineage, but not functional in the M lineage. When we examined its expression pattern, we found that *Tc1*::*gfp::sem-2* was not expressed in the M lineage at all (Figure 2.3A-C,E-G and data not shown), while its expression outside of the M lineage was largely unaffected (Figure 2.3D,H). We also forced the expression of *sem-2* cDNA in the M lineage of *n1343* animals using the M lineage specific *hlh-8* promoter and found that it rescued the missing egg-laying muscle phenotype in *sem-2(n1343)* animals, as determined by the reappearance of *egl-15::gfp* expression (6/6 lines). Together these observations demonstrate that the Tc1 insertion located in the 4.5kb intron specifically disrupts the M lineage expression of *sem-2*, and that *sem-2* is required within the M lineage for proper SM fate specification.

2.3.5 SEM-2/SoxC is sufficient to promote the SM fate

We then tested whether *sem-2* is sufficient to promote non-SM cells to adopt the SM fate. To this end, we used the *hlh-8* promoter to force *sem-2* cDNA expression in all the undifferentiated M lineage cells in wild-type animals. We first assayed for the effect of *sem-2* misexpression on the CC and BWM fates using an intrinsic CC specific reporter *CC::gfp* and a BWM specific reporter *myo-3p::rfp* (see Materials and Methods). Wild-type worms have 4 embryonically-derived and 2 M lineage-derived CCs (Figure 2.3I). Among animals carrying the *hlh-8p::sem-2* construct, 62.5% had no M-derived CCs (Figure 2.3J,K) and 14.1% had only one M-derived CC (n=64). The animals that lack M-derived CCs also lacked, on average, 10 of the 14 BWMs derived from the M lineage (n=9, Figure 2.3M,N). In contrast, 40.6% of the animals carrying *hlh-8p::sem-2* (n=64) had extra *hlh-8::gfp*-expressing SM-like cells and later, extra *egl-15::gfp*-expressing type I vulval muscle-like cells (Figure 2.3J). In most cases the vulval muscles born from animals misexpressing *sem-2* were not attached to the vulva properly or not even located in the vulval region, and were therefore not likely to be functional (Figure 2.3J,K). Taken together, the loss of M-derived BWM and CCs and the appearance of excessive SM- and vulval muscle-like cells are strongly suggestive of a transformation of some (if not all) of the M lineage cell types to SMs. Thus, *sem-2* is not only necessary, but also sufficient, to promote the SM fate.

Figure 2.3: *sem-2* is both required and sufficient in the M lineage for specifying the SM fate.

(A-H) Transgenic worms carrying *gfp::sem-2* (A-D) or *Tc1::gfp::sem-2* (E-H) at the 2-SM stage. *gfp::sem-2* (A-C), but not *Tc1::gfp::sem-2* (E-G), is expressed in the SM cells (one focal plane shown) labeled by *hlh-8::rfp* (B, F, with merged images shown in C and G). Both transgenes are expressed in vulval and hypodermal cells (D, H, ventral views). (I-N) wild-type (I, L) or transgenic worms carrying *hlh-8p::sem-2 cDNA* (J-K, M-N) that also express *cc::gfp*, which labels CC, and *egl-15::gfp*, which labels type I vulval muscles (I-K), as well as *myo-3::gfp*, which labels BWMs (L-N). Compared to wild-type animals, animals expressing *hlh-8p::sem-2* cDNA often have extra type I vulval muscles that are disorganized (enlargement boxes in J,K), lack M-derived CCs (J,K) and some M-derived BWMs (arrows in M-N indicate regions with missing M-derived BWMs). Arrowheads point to M-derived CCs. * denotes the location of the vulva.


gfp::
sem-2
A
2-SM
B
hlh-8::rfp
C
merge
D
Tc1::gfp
::sem-2
E
2-SM
F
hlh-8::rfp
G
merge
H
I
WT
cc::gfp; egl-15::gfp
J
hlh-8p::sem-2 cDNA
K
hlh-8p::sem-2 cDNA
L
myo-3::rfp
M
N

2.3.6 The M lineage specific expression of sem-2/SoxC is is controlled by elements in the 4.5 kb intron.

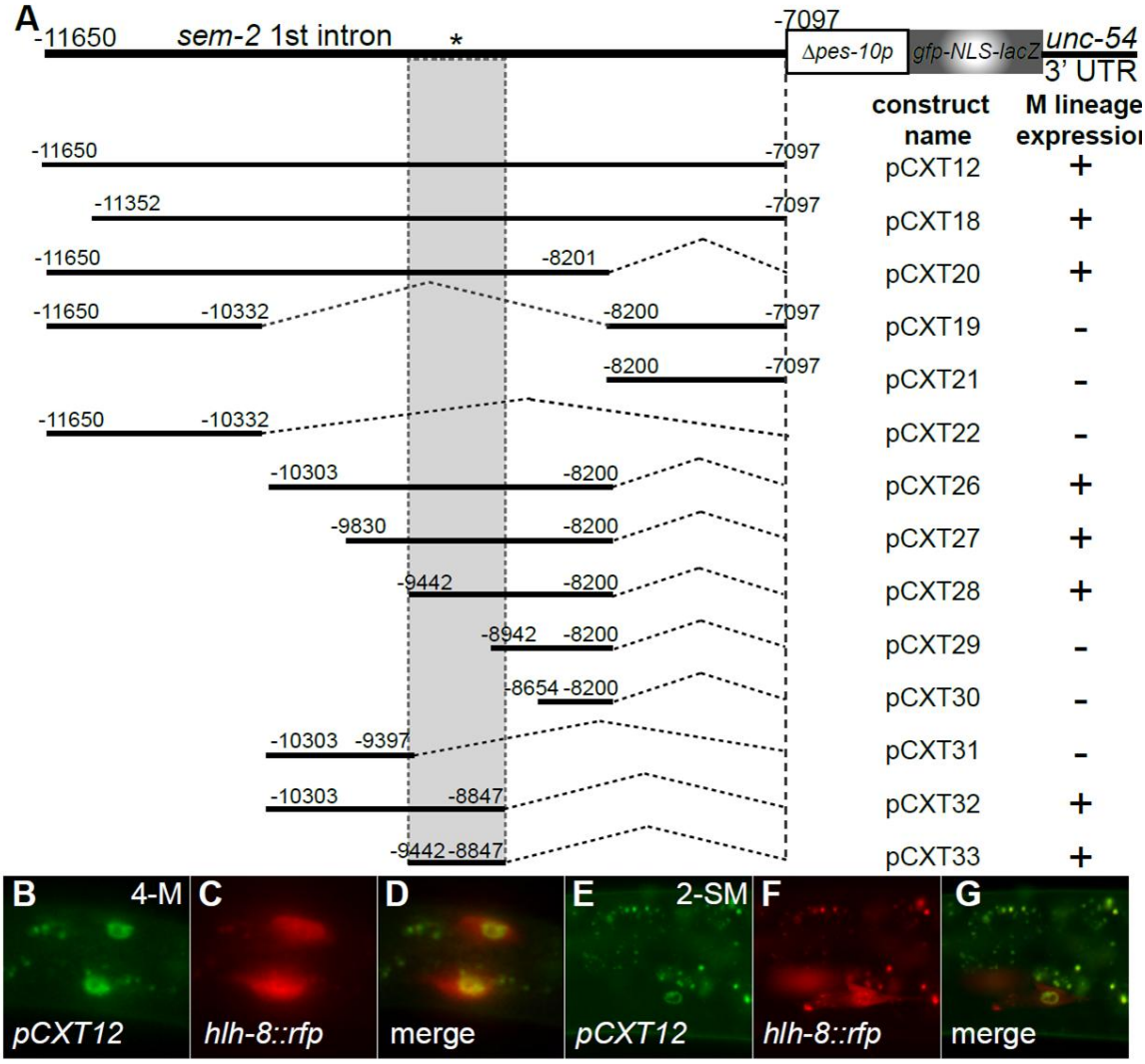
The M lineage specific function of *sem-2* in promoting the SM fate coupled with the presence of the Tc1 insertion in the 4.5kb intron and the loss of M lineage specific expression of *Tc1::gfp::sem-2* suggest that the 4.5kb intron may contain element(s) specifically required for *sem-2* expression in the M lineage. To test this hypothesis, we first placed the 4.5kb intron (Figure 2.2A) directly upstream of the *gfp* coding sequence, but observed no *gfp* expression in transgenic lines carrying the construct (2 lines, n>100). However, the 7kb intron of *sem-2* alone was capable of driving reporter expression in hypodermal, intestinal and vulval cells that express *sem-2* even in the *n1343* mutants (data not shown). Taking into account the two types of *sem-2* transcripts (Figure 2.2A), these observations suggest that the 7kb intron likely has the promoter responsible for the transcription of splice form 2 in hypodermal, intestinal and vulval cells, while the 4.5kb intron has no promoter activity on its own.

To further test whether the 4.5kb intron contains any M lineage specific enhancer(s), we placed the entire 4.5kb intron sequence upstream of the *Dpes-10* basal promoter and *gfp*, followed by the *unc-54* 3'UTR (pCXT12, Figure 2.4A). We observed M lineage specific expression of the reporter (6/8 lines, Figure 2.4B-G), suggesting the presence of M lineage specific enhancer element(s) within the 4.5kb intron. However, the *gfp* expression pattern from this reporter differs from that of the translational *gfp::sem-2* reporter described earlier; while GFP::SEM-2 was only found in the SM mother and the SM lineage cells,

transgenic lines carrying the pCXT12 reporter showed *gfp* expression in all the undifferentiated cells in the M lineage, including the early M lineage and the SM lineage (Figure 2.4B-G). We reasoned that either additional cis-element(s) are involved in restricting *sem-2* expression to specific cells within the M lineage, or the SEM-2 protein is unstable in the early M lineage. To this end, we tested 2.8kb sequences upstream of the 4.5kb intron, 2kb sequences of the 3' UTR, and the entire 7kb second intron by placing each of them in *cis* with the 4.5kb intron. None of them was able to restrict the expression of the *4.5kb intron::gfp* reporter in the early M lineage (Figure 2.4H). Despite this, our data demonstrate that the 4.5kb intron contains enhancer(s) positively involved in directing *sem-2* expression in the M lineage.

Figure 2.4: The 4.5kb intron of *sem-2* contains an M lineage enhancer element.

(A) Schematic of deletion series of the *sem*-2 4.5kb intron in the enhancer analysis. GFP expression was scored in F2 animals in at least 3 independent lines per construct. Expression is represented as + (consistent expression), - (no detectable expression). * denotes the location of Tc1 insertion in the *n1343* allele (also in panel H). (B-G) Examples showing lateral views (one focal plane) of worms expressing GFP driven by the pCXT12 reporter (B,E) in the M lineage (labeled by *hlh-8::rfp*, C,F) at the 4-M stage (B-D) and the 2-SM stage (E-G). (H) Constructs used to identify possible cis-negative elements in the *sem-2* locus. + indicates the presence of reporter expression.


A
-11650
sem-2 1st intron
*
-7097
Δpes-10p
gfp-NLS-lacZ
unc-54
3' UTR
construct name
M lineage expressio
-11650 -7097 pCXT12 +
-11352 -7097 pCXT18 +
-11650 -8201 pCXT20 +
-11650 -10332 -8200 -7097 pCXT19 -
-8200 -7097 pCXT21 -
-11650 -10332 pCXT22 -
-10303 -8200 pCXT26 +
-9830 -8200 pCXT27 +
-9442 -8200 pCXT28 +
-8942 -8200 pCXT29 -
-8654 -8200 pCXT30 -
-10303 -9397 pCXT31 -
-10303 -8847 pCXT32 +
-9442 -8847 pCXT33 +
B 4-M
pCXT12
C
hlh-8::rfp
D
merge
E 2-SM
pCXT12
F
hlh-8::rfp
G
merge

2.3.7 M lineage expression is under the direct control of Hox factors, MAB-5 and LIN-39, and their cofactor CEH-20

To identify the M lineage specific enhancer(s) within the 4.5kb intron, we generated a series of deletions in pCXT12 (Figure 2.4A), and found a 598bp region (-11650 to -7097) that is sufficient to drive reporter expression in the M lineage (Figure 2.4A). Alignment of this 598bp region in *C. elegans* and its homologous sequences in three related *Caenorhabditis* species, *C. briggsae, C. remanei* and *C. brenneri*, which are farther away to *C. elegans* than mouse to human, identified several blocks of highly conserved sequences (data not shown) that include a site “TGATATATCG” (Figure 2.5A). This site closely matches the consensus PBC/Hox binding sequence “TGATNNAT(G/T)(G/A)”, with “TGAT” being the PBC binding site and “AT(G/T)(G/A)” being the Hox binding site (Chan and Mann,1996;Mann and Affolter, 1998). Interestingly, the Tc1 transposon insertion site in the *sem-2(n1343)* mutants disrupts the putative PBC half site (Figure 2.5A). Furthermore, inserting Tc1 into the same location as in *n1343* mutants in the reporter construct pCXT33 completely blocked the M lineage expression of the reporter (pCXT173, 6/6 lines, n>100). We further tested the importance of the putative PBC/Hox binding site by making clustered mutations in each half site and testing the consequences of the mutations on the M lineage enhancer activity. Mutating the PBC half site (mut1) completely abolished the M lineage expression of the reporter (100%, n=45, Figure 2.5B). Mutations in the putative Hox binding site (mut2 and mut3) also led to the loss of the M lineage enhancer activity in all (100%, n=26, for mut2) or most (93.7%, n=63, for mut3) of

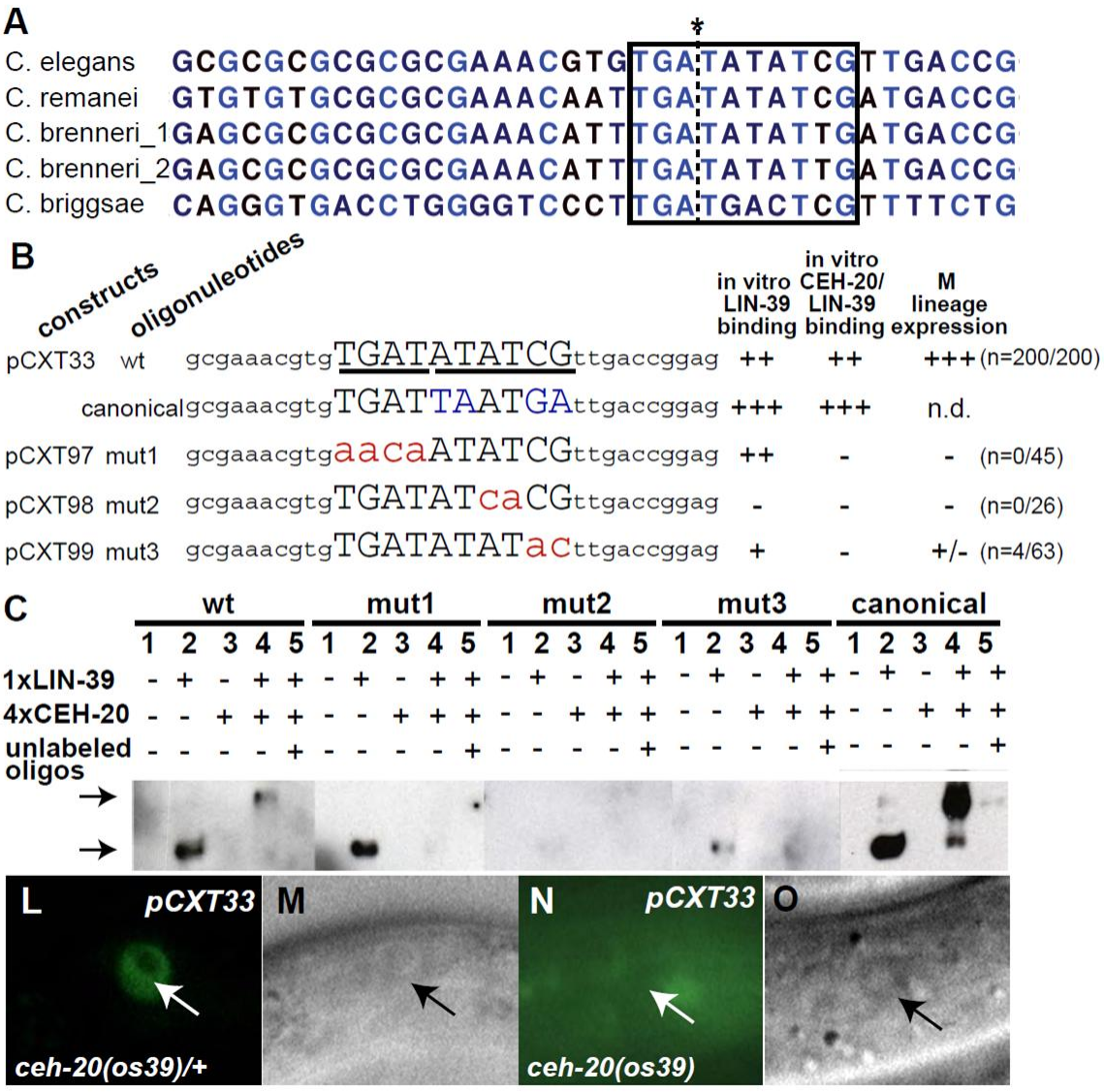
the transgenic animals (Figure 2.5B). Thus the putative PBC/Hox binding site, disrupted by the Tc1 insertion in *n1343* mutants, is required for *sem-2* M lineage activity.

Previous studies have found that two Hox factors, MAB-5 and LIN-39, and their cofactor, the PBC homolog CEH-20, play essential roles in M lineage development (Liu and Fire, 2000; Jiang et al., 2009). Both *mab-5* and *ceh-20* are expressed throughout the M lineage (Salser, 1995; Jiang et al., 2009), while *lin-39* is expressed in the SM lineage (Wagmaister et al., 2006). Furthermore, MAB-5 and LIN-39 together and CEH-20 are required for the proper specification and differentiation of M-derived CC, BWMs and SMs (Jiang et al., 2009). We therefore tested whether these three factors are required for the *sem-2* M lineage activity using the minimal pCXT33 reporter. We introduced pCXT33 into the double null mutant *lin-39(n1760) mab-5(e1239)* and a strong loss-of-function mutant *ceh-20(os39)* (Liu and Fire, 2000; Arata et al., 2006; Jiang et al., 2009). g*fp* reporter expression was detected in the M mesoblast in 98.3% (n=56) of *lin-39(n1760) mab-5(e1239)*/++ animals (data not shown) and 92.6% (n=54) of *ceh-20(os39)*/+ animals (Figure 2.5L, M). However, *gfp* expression was detected in the M mesoblast in only 4.9% (n=82) of *lin-39(n1760) mab-5(e1239)* animals and in 8.3% (n=72) of *ceh-20(os39)* animals (Figure 2.5N, O, and data not shown). Thus, the PBC/Hox factors MAB-5, LIN-39 and CEH-20 are necessary for the *sem-2* M lineage enhancer activity.

To test whether MAB-5, LIN-39 and CEH-20 may directly regulate *sem-2* expression in the M lineage, we used the Electrophoretic Mobility Shift Assay (EMSA) to test whether recombinant LIN-39 and CEH-20 proteins can cooperatively bind to the putative PBC/Hox binding site *in vitro*. We were not able to generate recombinant full-length MAB-5 proteins *in vitro* (Liu and Fire, 2000). As shown in Figure 2.5, LIN-39 alone, or together with CEH-20, binds oligonucleotides with a canonical ANTP/EXD composite site (Knoepfler et al., 1996) or oligonucleotides containing the putative PBC/Hox site in the *sem-2* M lineage enhancer (Figure 2.5B,C). In contrast, mutating the PBC half site (mut1) completely abolished the composite binding of LIN-39 and CEH-20 proteins without affecting LIN-39 binding alone (Figure 2.5C). Similarly, mutating the Hox half site (mut2) completely abolished the binding of LIN-39 alone or both LIN-39 and CEH-20. Consistent with the *in vivo* reporter assay result, mut3 significantly reduced, but did not completely abolish, the binding of LIN-39 alone or LIN-39 and CEH-20 together (Figure 2.5C). The composite binding of LIN-39 and CEH-20 to the putative PBC/Hox binding site can also be competed away using excess of unlabeled oligonucleotides containing the wild-type binding site (Figure 2.5C). Taken together, our data suggest that *sem-2* is directly regulated by the Hox factors MAB-5 and LIN-39 and their cofactor CEH-20 in the M lineage. Consistent with our finding, ChIP-seq experiments by modENCODE showed that LIN-39 directly binds to this Hox-PBC site *in vivo* (http://intermine.modencode.org).

Figure 2.5: *sem-2* is a direct target of Hox/PBC factors in the M lineage.

(A) Alignment of the conserved intronic regions that include the putative Hox/PBC binding site (box region) from *C. elegans*, *C. remanei*, *C. brenneri* and *C. briggsae* using clustalW. *C. brenneri* has two copies of the *sem-2* gene. * marks the Tc1 insertion site in *sem-2(n1343)*. Blue color indicates identity among the four species. (B) Summary of *in vivo* transgenic reporter assays and *in vitro* EMSA results. The putative PBC (red) and Hox (blue) binding sites are in capital letters. Mutated sites are in lower case. +++: strong binding or robust reporter expression, ++: moderate binding, +/-: faint expression in 4 of 63 animals examined, -: no binding or no detectable reporter expression. (C) EMSA using indicated oligonucleotides, purified LIN-39 and CEH-20 protein, and un-labeled competitor wild-type oligonucleotides in 2000-fold excess. The top and bottom arrows indicate the sizes of LIN-39/CEH-20/DNA and LIN-39/DNA complexes, respectively. (L-O) pCXT33 expression in *ceh-20(os39)/+* (L,M) and *ceh-20(os39)* (N,O) animals. (L,N) GFP, (M,O) DIC. Arrow: M mesoblast.

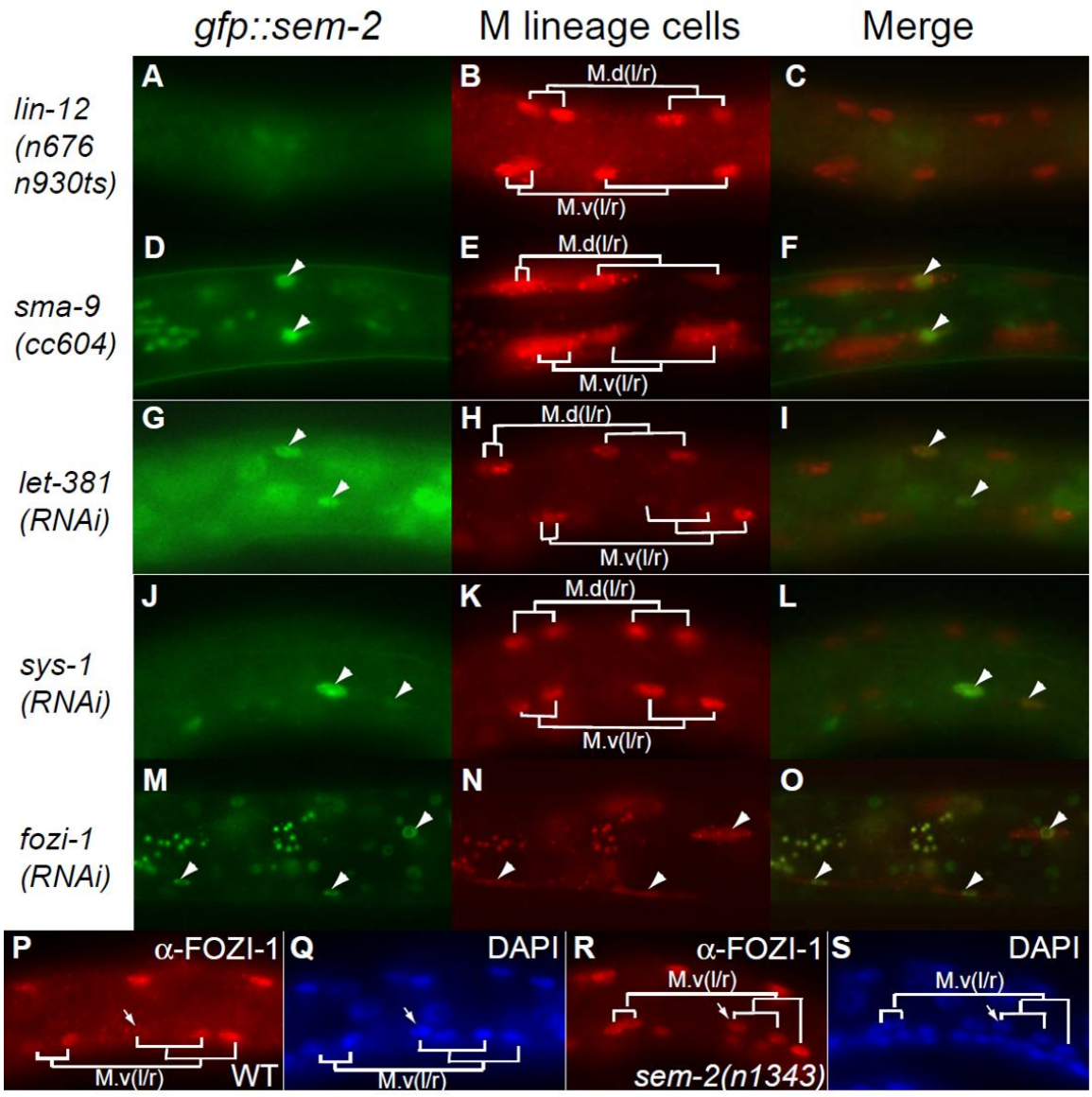

A
C. elegans GCGCGCGCGCGCGAAACGTGTGATATATCGTTGACCG
C. remanei GTGTGTGCGCGCGAAACAATTGATATATCGATGACCG
C. brenneri_1 GAGCGCGCGCGCGAAACATTTGATATATTGATGACCG
C. brenneri_2 GAGCGCGCGCGCGAAACATTTGATATATTGATGACCG
C. briggsae CAGGGTGACCTGGGGTCCCTTGATGACTCGTTTTCTG
B
constructs
oligonuleotides
in vitro LIN-39 binding
in vitro CEH-20/ LIN-39 binding
M lineage expression
pCXT33 wt gcgaaacgtgTGATATATCGttgaccggag ++ ++ +++ (n=200/200)
canonical gcgaaacgtgTGATTAATGAttgaccggag +++ +++ n.d.
pCXT97 mut1 gcgaaacgtgaacaATATCGttgaccggag ++ - - (n=0/45)
pCXT98 mut2 gcgaaacgtgTGATATcaCGttgaccggag - - - (n=0/26)
pCXT99 mut3 gcgaaacgtgTGATATATacttgaccggag + - +/- (n=4/63)
C
wt
mut1
mut2
mut3
canonical
1 2 3 4 5
1xLIN-39
4xCEH-20
unlabeled oligos
L pCXT33
ceh-20(os39)/+
M
N pCXT33
ceh-20(os39)
O

2.3.8 SEM-2/SoxC acts downstream of signaling pathways required for proper SM fate specification

Previous studies have shown that LIN-12/Notch signaling is required for promoting the SM fate on the ventral side, while the antagonism of TGF-β signaling by SMA-9 is required for the CC fate on the dorsal side, of the M lineage (Greenwald et al., 1983; Foehr et al., 2006; Foehr and Liu, 2008). We therefore tested whether SM fate and *sem-2* expression in the M lineage is under the control of these dorsoventral patterning mechanisms. As shown in Figure 2.6A-C, no *gfp::sem-2* expression was observed in M.v(l/r)pa cells at the 16-M stage (n=44) in *unc-32(e189) lin-12(n676n930ts)* animals at the restrictive temperature, in which the SMs are transformed to CCs (Greenwald et al, 1983; Foehr and Liu, 2008). In contrast, *gfp::sem-2* expression was detected in both M.d(l/r)pa and M.v(l/r)pa as well as their descendants at the 16-M stage (n=18, Figure 2.6D-F) and the 18-M stage (n=4) respectively in *sma-9(cc604)* mutants, which show a M-derived CC to SM fate transformation (Foehr et al., 2006). Furthermore, *sem-2* is required for specifying the dorsal SMs in *sma-9(cc604)* mutants, as both the endogenous and the ectopic SMs in *sma-9(cc604)* mutants were lost upon *sem-2(RNAi)* treatment (97.3%, n=75) or in *sem-2(n1343); sma-9(cc604)* double mutants (100%, n>200). Thus, the M lineage expression of *sem-2* is downstream of both LIN-12 and SMA-9.

Figure 2.6: The M lineage expression of *sem-2* is under the control of multiple signaling pathways and transcription factors.

(A-O) *sem-2* M lineage expression in *lin-12(n676n930ts)* animals at the restrictive temperature (A-C), *sma-9(cc604)* (D-F), *let-381(RNAi)* (G-I), *sys-1(RNAi)* (J-L) and *fozi-1(RNAi)* (M-O) animals. Arrowheads indicate cells with *gfp::sem-2* expression. The animal shown in (M-O) is at the 2-SM stage. (P-S) *fozi-1* is ectopically expressed in M.vlpaa in *sem-2(n1343)* animals (R-S) compared to wild-type animals (P-Q), as shown by α-FOZI-1 immunostaining (P, R) and the corresponding DAPI staining images (Q, S). Arrows point to M.vlpaa. Similar expression is also seen in the equivalent cells on the right side for A-L and P-S.


gfp::sem-2
M lineage cells
Merge
lin-12 (n676 n930ts)
sma-9 (cc604)
let-381 (RNAi)
sys-1 (RNAi)
fozi-1 (RNAi)
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
M.d(l/r)
M.v(l/r)
P
α-FOZI-1
WT
Q
DAPI
R
α-FOZI-1
sem-2(n1343)
S
DAPI

We have recently shown that the FoxF/C forkhead transcription factor LET-381 also functions downstream of SMA-9 and LIN-12 to promote CC fate specification (Amin et al., 2010). *let-381* is expressed in M.d(l/r)pa and their mothers and *let-381(RNAi)* leads to the fate transformation of M-derived CCs to the SM mother cells (Amin et al., 2010). In *let-381(RNAi)* animals, *gfp::sem-2* was detected in both M.d(l/r)pa and M.v(l/r)pa as well as their descendants (Figure 2.6G-I, n=23/45). Furthermore, no SMs were produced in *sem-2(RNAi) let-381(RNAi)* animals (94.1%, n=118), suggesting that *sem-2* functions downstream of *let-381*. Thus the exclusive ventral expression of *sem-2* in the M lineage is due to the negative regulation of LET-381, whose own expression is repressed by LIN-12 on the ventral side and activated by SMA-9 on the dorsal side (see Figure 2.7).

Along the anterior-posterior axis, the Wnt/β-catenin asymmetry pathway is required for proper M lineage fate specification at the 16-M stage (Amin et al., 2009). Specifically, SYS-1/β-catenin is enriched in the nuclei of posterior cells and mutations of *sys-1* lead to a posterior-to-anterior fate transformation and the production of extra SMs and CCs, while POP-1/TCF has a reciprocal localization and loss-of-function phenotypes (Amin et al., 2009). Because *sem-2* expression is first detected only in M.v(l/r)pa cells but not their posterior sister cells M.v(l/r)pp at the 16-M stage, we looked to see if *sem-2* expression in the M lineage is under the control of the Wnt/β-catenin asymmetry pathway. As shown in Figure 2.6J-L, *gfp::sem-2* was expressed in both M.v(l/r)pa and M.v(l/r)pp (33.3%, n=26) and all

the endogenous and ectopic SMs (85.7%, n=14) in *sys-1(RNAi)* animals. *sem-2* is also required for both the endogenous and the ectopic SMs produced in *sys-1(q544)* animals, as *sem-2(RNAi) sys-1(q544)* animals produced no SMs at all (100%, n=20). Thus *sem-2* acts downstream of, and is negatively regulated by, SYS-1.

2.3.9 sem-2 exhibits mutually repressive interactions with fozi-1 and hlh-1.

Previous studies have shown that the MyoD homolog HLH-1 functions redundantly with the zinc finger protein FOZI-1 to specify M-derived BWMs and CCs while repressing the SM fate (Harfe et al., 1998a; Amin et al., 2007). We found that *gfp::sem-2* was expressed in both the endogenous and the ectopic SMs in *fozi-1(RNAi)* animals (98.6%, n=70, Figure 2.6M-O) and *hlh-1(RNAi)* animals (100%, n=25). Furthermore, no SMs were produced by *sem-2(RNAi) fozi-1(cc609)* animals (100%, n=71) or *sem-2(RNAi) hlh-1(cc561ts)* animals at the restrictive temperature (100%, n=30). These observations suggest that *sem-2* expression is repressed in the M-derived BWMs and CCs by HLH-1 and FOZI-1.

Because the SMs are transformed to BWMs in *sem-2(n13430)* mutants, we also examined whether *fozi-1* and *hlh-1* are expressed in the ectopic BWMs in *n1343* animals. As shown in Figure 2.6P-S, *fozi-1*, which is transiently expressed in all M-derived BWMs prior to their differentiation in wild-type animals (Figure 2.6P-Q), is expressed in the ectopic BWMs in *sem-2(n1343)* animals (Figure 2.6R-S). Similar results were also obtained for *hlh-1* expression in *sem-*

2(n1343) animals (data not shown). Since the expression of *sem-2* in the SM mother cells (M.v(l/r)pa) overlaps with those of *hlh-1* and *fozi-1*, the above results suggest that the BWM-specifying factors FOZI-1 and HLH-1 and the SM-specifying factor SEM-2 mutually repress each other's expression to maintain their proper expression pattern in the respective daughter cells derived from the SM mothers.

2.4 DISCUSSION

2.4.1 SEM-2/SoxC acts as a switch in a binary fate decision to promote a proliferative fate over a differentiated muscle fate

We have demonstrated that the single SoxC protein SEM-2 is both necessary and sufficient to specify the SM fate in the *C. elegans* postembryonic mesoderm. SMs are precursors that have the potential to proliferate and then differentiate into 16 non-striated egg-laying uterine and vulval muscles. Disrupting *sem-2* expression specifically in the M lineage leads to a transformation of SMs to cells that differentiate into striated BWMs. *sem-2* is specifically expressed in Mv(l/r)pa, a bi-potent precursor cell that asymmetrically divides to give rise to a SM and a BWM. *sem-2* expression is retained in the SM, perdures in its descendants that remain proliferative and ceases when these cells switch from a proliferation state and differentiate into mature egg-laying muscles (Figure 2.2R). Forced over-expression of *sem-2* throughout the M lineage led to the loss of M-derived BWMs and CCs and the gain of extra SMs

(Figure 2.3J-N). These experiments demonstrate that SEM-2 is both necessary and sufficient for promoting the egg-laying muscle precursor SM fate while antagonizing the *sem-2*-expressing cells from differentiating into BWMs. Consistent with this, ectopic *sem-2* expression was detected in mutant animals with extra SMs and removing *sem-2* in these animals blocked the formation of both ectopic and normal SMs.

sem-2 is also broadly expressed during embryogenesis and essential for embryonic development. *sem-2* null mutant embryos were arrested at the 3-fold stage. Time-lapse movies showed that this arrest is preceded by delays in hypodermal migration.and elongation. Besides the embryonic phenotype, RNAi knockdown of *sem-2* postembryonically caused a multivulval (Muv) phenotype, which may result from a fate specification defect or a proliferation defect in the vulval precursor cells. Further characterization of the embryonic and vulval phenotypes of *sem-2* mutants will help shed light on how *sem-2* functions in these processes.

2.4.2 A model for the specification of the non-striated muscle precursors, the SMs

Our findings provide an example of how lineage information and positional information are integrated to activate the cell type specific expression of a cell fate determinant, in this case, *sem-2*, in the specification of the sex myoblast cells. The M lineage expression of *sem-2* is controlled by a *cis*-acting M lineage enhancer located in the 4.5kb intron (Figure 2.4A). Disruption of this site by the

transposon Tc1 in *n1343* animals specifically disrupted the expression and function of *sem-2* in the M lineage without affecting its expression in other cells. We showed that the Hox factors MAB-5 and LIN-39 and their cofactor, the PBC protein CEH-20, directly bind this site to activate *sem-2* expression in the M lineage (Figure 2.5B,C). However, these three genes are also expressed in other cell types in *C. elegans* (Costa et al., 1988; Wagmaister et al., 2006; Jiang et al., 2009). In fact, it has been shown that MAB-5/CEH-20 and/or LIN-39/CEH-20 complexes directly activate the expression of *hlh-8* throughout the M lineage before terminal differentiation (Liu and Fire, 2000); of *egl-18, elt-6, eff-1* and *lag-2* in the developing vulva (Koh et al., 2002;Shemer and Podbilewicz, 2002;Takacs-Vellai et al., 2007), and of *egl-1* in the P11 lineage and the VC neurons (Liu et al., 2006;Potts et al., 2009). A Hox independent role of CEH-20 has also been found in the activation of *mls-2* in the early M lineage (Jiang et al., 2008). Thus, additional factors must cooperate with Hox/PBC proteins to refine the cell type specific expression of Hox/PBC targets.

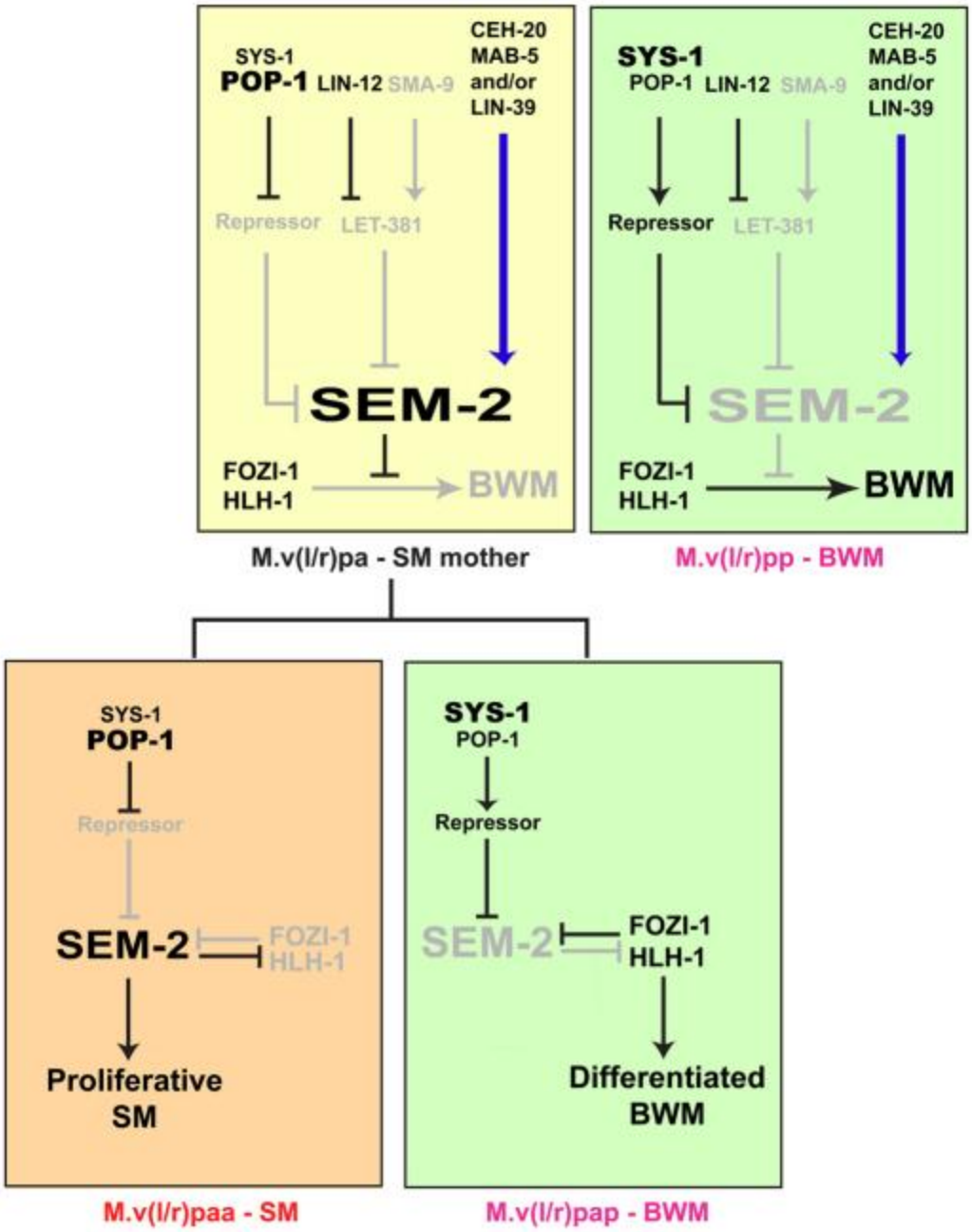
Our results demonstrate that, in addition to the Hox factors and CEH-20, the integration of dorsal-ventral (D/V) and anterior-posterior (A/P) positional information is critical in dictating the specific localization of *sem-2* within the M lineage. Previous studies have shown that LIN-12/Notch signaling and SMA-9, which antagonizes the Sma/Mab TGF-β signaling pathway, work independently to control D/V patterning of the M lineage and that they both regulate the expression of the FoxF/C transcription factor LET-381 for its dorsal M lineage expression (Foehr et al., 2006; Foehr and Liu, 2008; Amin et al., 2010). We

showed that the proper pattern of *sem-2* expression is due to the presence of LIN-12/Notch signaling and the absence of SMA-9 along the D/V axis. Furthermore, *sem-2* functions downstream of *let-381* and is negatively regulated by *let-381* in the dorsal M lineage (Figure 2.6G-I). Currently it is not clear whether this negative regulation of *sem-2* by *let-381* is direct or not.

Along the A/P axis at the 16-M stage, POP-1/TCF functions as a repressor in M.v(l/r)pa due to low nuclear levels of SYS-1/β-catenin and is converted to an activator in M.v(l/r)pp due to high nuclear levels of SYS-1/β-catenin (Kidd et al., 2005;Liu et al., 2008; Amin et al., 2009). We showed that *sys-1(RNAi)* caused ectopic expression of *gfp::sem-2* and ectopic production of SMs (Figure 2.6J-L). Since *sem-2* is expressed in M.v(l/r)pa where POP-1 acts as a repressor, these results suggest that POP-1 must activate the expression of a repressor of *sem-2* in M.v(l/r)pa to restrict *sem-2* expression along the anterior-posterior axis in the M lineage. We hypothesize that an additional factor(s) must be present either in M.v(l/r)aa to repress *sem-2* expression or in M.v(l/r)pa to promote *sem-2* expression in order to ensure specific expression of *sem-2* in M.v(l/r)pa (see model in Figure 2.7).

Figure 2.7: A Model for *sem-2* regulation and SM fate specification.

A proposed model on how *sem-2* expression in the M lineage is regulated (see discussion). * denotes a putative transcriptional repressor of *sem-2*. Thick blue lines represent direct regulatory relationships, black lines represent relationships based solely on genetic data, and do not distinguish between direct and indirect. Gray lines and text indicate a lack of expression in the indicated cell.


SYS-1
POP-1
LIN-12
SMA-9
CEH-20
MAB-5
and/or
LIN-39
Repressor
LET-381
SEM-2
FOZI-1
HLH-1
BWM
M.v(l/r)pa - SM mother
M.v(l/r)pp - BWM
Proliferative
SM
Differentiated
BWM
M.v(l/r)paa - SM
M.v(l/r)pap - BWM

Immediately after the SM mother cell M.v(l/r)pa divides, both its daughter cells express *gfp::sem-2* (Figure 2.2I-K). However, *gfp::sem-2* perdures in the anterior SM lineage, but not in the posterior BWMs. *hlh-1* and *fozi-1* exhibit a reciprocal expression pattern: they are turned off in the SMs, but remain expressed in the BWMs (Harfe et al., 1998a; Amin et al., 2007). The initiation of the asymmetric expression of *sem-2* vs. *hlh-1* and *fozi-1* may be due to the nuclear POP-1 and SYS-1 asymmetry along the anterior-posterior axis (Amin et al., 2009). Once initiated, *sem-2* and *hlh-1* and *fozi-1* appear to mutually repress each other to maintain their proper expression, as *sem-2* is expressed in all the ectopic SMs in *hlh-1* and *fozi-1* mutant or RNAi animals, while *fozi-1* and *hlh-1* are expressed in the ectopic BWMs in *sem-2(n1343)* animals.

2.4.3 An evolutionarily conserved role of SoxC family members in cell fate specification and cell proliferation

SEM-2 is a member of the SoxC subfamily, which includes Sox14 in Drosophila and Sox4, 11 and 12 in vertebrates (Bowles et al., 2000). Drosophila Sox14 exhibits a dynamic expression pattern during development and regulates dendrite severing and other unknown processes essential for fly metamorphosis (Sparkes et al., 2001;Kirilly et al., 2009;Ritter and Beckstead, 2010). Functional studies on vertebrate SoxC members have shown that SoxC proteins play essential roles in multiple lineages such as oligodendrocytes, B lymphocytes, osteoblasts and others to regulate cell fate specification, cell differentiation and cell survival (Penzo-Mendez, 2009). In humans, Sox4 overexpression has often

been detected in cases of prostate cancer (Dhanasekaran et al., 2001;Ernst et al., 2002;Lapointe et al., 2004;Liu et al., 2006;Luo et al., 2001;Magee et al., 2001;Rhodes et al., 2002;Welsh et al., 2001). Sox4 is also overexpressed in many other types of cancers, including leukemias (Andersson et al., 2007), melanomas (Talantov et al., 2005), glioblastomas (Sun et al., 2006), medulloblastomas (Lee et al., 2002), bladder cancer (Aaboe et al., 2006) and lung cancer (Friedman et al., 2004). Similarly, Sox11 is highly expressed in medulloblastomas (Lee et al., 2002), gliomas (Weigle et al., 2005); non-B cell lymphomas (Wang et al., 2008) and epithelial ovarian tumors (Brennan et al., 2009). However, the underlying mechanism by which the Sox proteins may contribute to cancer is not known. We have found that *sem-2* is specifically expressed in the proliferating SM mother cells, the SMs and their descendants prior to terminal differentiation (Figure 2.2R). Furthermore, increasing *sem-2* level throughout the M lineage is sufficient to transform other M lineage cells into proliferating SM-like cells, even though they express normal levels of CC- and BWM-specifying and differentiating factors. Thus it appears that increasing *sem-2* level is sufficient to tip the balance between proliferative and differentiative factors and allows for proliferation. This pro-proliferation function of SEM-2 may be a conserved role of the group C Sox proteins.

Hox and PBC proteins have also been implicated in tumorigenesis. There are numerous examples of Hox gene overexpression in various tumor types (Argiropoulos and Humphries, 2007;Shah and Sukumar, 2010). Given our finding that *sem-2* expression in the M lineage is directly activated by Hox-PBC proteins

and that there is overexpression of Sox4 and Sox11 in a variety of cancers, it is possible that the oncogenic activity of Hox genes in some cases is due to their direct activation of SoxC gene expression.

2.5 ACKNOWLEDGEMENT

We thank the *C. elegans* Genetics Center, the knockout consortium, and Yuji Kohara for strains and yk clones; Nirav Amin for examining *sem-2* expression in *let-381(RNAi)* animals; Alex Beatty and Rich McCloskey for help with preparing embryos for the time lapse movies; Nirav Amin and Jared Hale for helpful discussions and critical comments on the manuscript. This work was supported by NIH R01 GM066953 (to J.L.).

Chapter 3: The RGM protein DRAG-1 positively regulates a BMP-like signaling pathway in *Caenorhabditis elegans*[2]

3.1 INTRODUCTION

Bone morphogenetic proteins (BMPs) belong to the transforming growth factor-β (TGFβ) superfamily of ligands and the BMP signaling pathway plays roles in multiple developmental and homeostatic processes (Wu and Hill, 2009). Malfunction of the pathway causes many somatic and hereditary disorders in humans, including cardiovascular diseases and cancer (Gordon and Blobe, 2008; Massagué, 2008). All BMP family members share a common mode of signal transduction (Shi and Massague, 2003). Upon binding of the BMP ligands, the type II receptor phosphorylates the type I receptor, which then phosphorylates the receptor-regulated Smads (R-Smads). These activated R-Smads then complex with common-mediator Smads (co-Smads) and enter the nucleus where the complex participates in regulation of downstream gene expression. Multiple levels of regulation, including extracellular regulation, ensure proper spatiotemporal control of BMP signaling in the right cellular context (Massague and Chen, 2000; Balemans and Van Hul, 2002; Moustakas and Heldin, 2009; Umulis et al., 2009). In particular, a family of repulsive guidance molecules (RGMs) has recently been found to bind different BMPs with high

[2] This chapter was published in *Development* (Tian C, Sen D, Shi H, Foehr ML, Plavskin Y, Vatamaniuk OK, Liu J. Development. 2010 Jul;137(14):2375-84.) and is reprinted with permission. Mapping was performed by Y. Plavskin, M.L. Foehr, and D. Sen, molecular cloning was performed by H. Shi, fractionation experiment was supervised by O.K.Vatamaniuk.

affinity and acts as BMP co-receptors in tissue culture experiments (Samad et al., 2005; Babitt et al., 2005, 2006).

RGM family proteins are GPI (glycophosphatidylinositol)-linked membrane-associated proteins (Corradini et al., 2009; Severyn et al., 2009). Vertebrates have four RGMs: RGMa, RGMb (DRAGON), RGMc (haemojuvelin or HJV) and RGMd (Camus and Lambert, 2007). RGMa, RGMb/DRAGON and RGMc/haemojuvelin can bind selected BMP molecules as well as type I and type II BMP receptors to enhance BMP signaling in tissue cultures (Babitt et al., 2005, 2006; Samad et al., 2005; Xia et al., 2008, 2010; Andriopoulos et al., 2009). Over-expression of mouse RGMb/DRAGON in Xenopus embryos can augment Smad1-induced mesodermal and endodermal marker expression (Samad et al., 2005). Several lines of *in vivo* evidence suggest that RGMc/haemojuvelin acts as the co-receptor for BMP-6 in regulating iron metabolism in mice (Babitt et al., 2007; Xia et al., 2008; Andriopoulos et al., 2009). However, there is as yet no evidence showing that RGMa and RGMb/DRAGON are required for modulating BMP signaling *in vivo*.

In this study, we describe the identification and functional characterization of the sole *C. elegans* RGM homolog DRAG-1 in modulating a BMP-like signaling pathway. There are two TGFβ-related pathways in *C. elegans*, a TGFβ-like pathway that controls dauer formation and a BMP-like Sma/Mab pathway that regulates body size, male tail formation, mesoderm development and innate immunity (Savage-Dunn, 2005; Foehr et al., 2006; Moustakas and Heldin, 2009;

Partridge et al., 2010). The ligand of the Sma/Mab pathway is a BMP-like molecule DBL-1 (Morita et al., 1999; Suzuki et al., 1999). Additional members of this pathway include the type I receptor SMA-6 (Krishna et al., 1999), type II receptor DAF-4 (Estevez et al., 1993), and Smads SMA-2, SMA-3 and SMA-4 (Savage et al., 1996). Loss-of-function mutations in any of these pathway members cause small body size and male tail sensory ray formation defects.

The Sma/Mab pathway also plays a role in patterning the *C. elegans* postembryonic mesoderm. The hermaphrodite postembryonic mesodermal M lineage arises from a single pluripotent precursor cell, the M mesoblast. During larval development, the M mesoblast divides to produce a dorsal lineage that gives rise to macrophage-like coelomocytes (CCs) and striated bodywall muscles (BWMs), and a ventral lineage that produces BWMs and the sex muscle precursor cells, the sex myoblasts (SMs) (Sulston and Horvitz, 1977; Figure 3.1C). We have previously shown that mutations in the *schnurri* homolog *sma-9* lead to a dorsal to ventral fate transformation in the M lineage (Foehr et al., 2006; Foehr and Liu, 2008). Furthermore, mutations in the core components of the Sma/Mab pathway suppress the dorsoventral patterning defects of *sma-9* mutants, suggesting that SMA-9 functions by antagonizing Sma/Mab signaling to pattern the postembryonic mesoderm along the dorsoventral axis.

Among suppressors of the *sma-9* phenotype, we detected a mutation in a novel locus that we have named *drag-1*. *drag-1* mutants exhibit only a subset of phenotypes seen in the Sma/Mab pathway mutants. We show that *drag-1*

encodes the sole RGM homolog in *C. elegans*. We also show that DRAG-1 is a membrane-localized protein that functions at the ligand-receptor level in the Sma/Mab pathway to regulate body size and mesoderm patterning. Using a novel Sma/Mab responsive reporter, we demonstrate that DRAG-1 positively modulates Sma/Mab signaling. Our work establishes a direct link between RGM proteins and BMP signaling *in vivo* and provides a simple and genetically tractable system for mechanistic studies on RGM protein regulation of BMP pathways *in vivo*.

3.2 MATERIAL AND METHODS

3.2.1 C. elegans strains

Strains were kept at normal conditions, as described by Brenner (Brenner, 1974). Analyses were performed at 20°C, unless otherwise noted. The following mutations and integrated transgenes were used: Linkage group II (LGII): *sma-6(jj6)*; LGIII: *daf-7(m62)*, *daf-7(e1372)*, *daf-4(m72)*, *sma-3(jj3)*, *lon-1(e185)*, *ccIs4438(intrinsic CC::gfp);* LGIV: *daf-1(m40)*, *daf-1(m213)*; LGV: *dbl-1(wk70)*; LGX: *lon-2(e678), sma-9(cc604), sma-9(cc606), sma-9(wk5)*. Other strains used: VF1: *unc119(ed3); gfEx[phmt-1p::gfp, unc-119(+)]* (Schwartz et al., 2010); VF12.3: *hmt-1(gk161)* III; *gfIs1[hmt-1::gfp, unc-119(+)].*

3.2.2 Molecular analysis of drag-1

The molecular lesion of *jj4* was identified by sequencing PCR fragments spanning the entire genomic region of Y71G12B.16.

To determine the *drag-1* cDNA sequence, the ProQuest cDNA library (Invitrogen) was used for PCR amplification using primers ForwLongpPC-86 (corresponding to vector sequences—5' TATAACGCGTTTGGAATCACTACAGGGATGTTTAATACCAC 3') and DS147 (complementary to sequences spanning the predicted stop codon of *drag-1*—5' TCAGCATAACAATGATAAAAGAGCAAAAAAAAAG 3'). The PCR fragments were cloned into the pCR-Blunt II-TOPO vector (Invitrogen). Three of the 10 resulting clones sequenced contained the entire Y71G12B.16 ORF as well as additional 5' UTR sequences. One clone (pDS9.10) has 32bp of 5' UTR sequences immediately upstream of the ATG (Y71G12B.16b, Genbank: HM154524), while the other two (pDS9.5 and pDS9.6) contain 60bp 5' UTR sequences (Y71G12B.16a, Genbank: HM154523).

3.2.3 Plasmid constructs and transgenic lines

drag-1 reporter constructs

4kb of the *drag-1* upstream sequence (-3977 to -1), sequences from immediately downstream of the ATG till the end of the first intron of *drag-1* (4 to 1123), the entire *drag-1* protein coding region and 1 kb of the *drag-1* downstream sequences were PCR amplified using N2 genomic DNA as template. The PCR

products were used to generate the following reporter constructs for analyzing the expression pattern of *drag-1*:

pDS14: 4kb *drag-1p::4xNLS::gfp::unc-54* 3' UTR;

pDS15: 4kb *drag-1p::drag-1 1st intron::4xNLS::gfp::unc-54* 3' UTR (Transgenic animals carrying either pDS14 or pDS15 gave identical *gfp* expression patterns, data not shown).

pDS27: 4kb *drag-1p::drag-1 genomic ORF::*1kb *drag-1* 3' UTR

pCXT15: 4kb *drag-1p::drag-1 1st intron::drag-1 cDNA::unc-54* 3'UTR

pJKL849: 4kb *drag-1p::drag-1 genomic ORF::gfp::*1kb *drag-1* 3' UTR. GFP from the Fire lab vector pPD95.75 (http://www.addgene.org) was inserted between aa395 and aa396 of the *drag-1* coding region, just prior to the cleavage site of the putative C-terminal pro-peptide.

Constructs for tissue specific expression of drag-1

pJKL928: *myo-2p::drag-1* cDNA::*unc-54* 3' UTR

pJKL931: *elt-2p::drag-1* cDNA::*unc-54* 3' UTR

pJKL933: *elt-3p::drag-1* cDNA::*unc-54* 3' UTR

pJKL935: *rol-6p::drag-1* cDNA::*unc-54* 3' UTR

pDS20: *hlh-8p::drag-1* cDNA::*unc-54* 3' UTR

pCXT148: *hlh-8p::drag-1* cDNA::1 kb *drag-1* 3' UTR

Constructs for structure-function analysis of *drag-1* (derived from pJKL849)

pCXT92 (DelC::GFP): *drag-1::gfp* with the putative C-terminal pro-peptide (aa387-408) deleted

pCXT115 (DelN::GFP): *drag-1::gfp* with the N-terminal signal peptide (aa2-23) deleted

pCXT116 (DelNC::GFP): *drag-1::gfp* with both the N-terminal signal peptide (aa2-23) and the C-terminal pro-peptide (aa387-408) sequences deleted

pCXT94 (LIN-12TM::GFP): *drag-1::gfp* with the C-terminal pro-peptide (aa387-408) replaced by the LIN-12 transmembrane domain (aa907-934 of LIN-12, Yochem et al., 1988)

Other reporters:

pJKL840: *sma-6p::nls::rfp::lacZ::unc-54 3' UTR* (containing 3kb of the *sma-6* promoter (-3094 to -1))

Generating transgenic animals

Transgenic animals were generated using the plasmids pRF4 (Mello et al., 1991), pC1 (*pha-1* rescuing construct, Granato et al., 1994), pJKL449 (*myo-2p::gfp::unc-54* 3' UTR, Jiang et al., 2009) or LiuFD61(*mec-7p::mRFP*, Amin et al., 2009) as markers.

3.2.4 Reverse Transcription-PCR (RT-PCR)

Total RNA was extracted from mixed stage wild-type and *drag-1(jj4)* worms using a RNA extraction kit (Qiagen). Reverse transcription was performed using the Superscript III First Strand Synthesize kit (Invitrogen) following the manufacturer's instructions. Primers used for amplifying the wild-type and *jj4* cDNAs are CXT49 (annealing to exon 1, forward) and CXT50 (annealing to exon 3, reverse). The resulting PCR fragments were purified and sequenced using primer CXT67.

3.2.5 Body size measurement

Hermaphrodite animals at the gravid adult stage were collected and treated with hypochlorite. The resulting embryos were allowed to hatch in M9 buffer at 16°C. Synchronized L1s were plated onto NGM plates and allowed to grow for 24h, 48h, 72h or 96h respectively before they were washed off the plates, treated with 0.3% sodium azide, and mounted onto 2% agarose pads. Images of the worms were taken on a Leica DMRA2 compound microscope with a Hamamatsu Orca-ER Camera using the Openlab software (Improvision). Body length of each animal was measured using Openlab software. Subsequent statistics analyses were performed using Microsoft Excel.

3.2.6 Dauer formation assays

Frequency of dauer formation was assessed under non-dauer-inducing conditions as described (Vowels and Thomas, 1992). Five to 10 adult

hermaphrodites were placed on a 6cm NGM plate at the test temperature. After a short period of egg-lay (less than 12 hours at 25°C), the parents were removed from the plates and their progeny were allowed to develop at a given temperature. When the non-dauer worms became young adults, the numbers of dauers and non-dauers on each plate were counted.

3.2.7 Immunofluorescence staining

Animal fixation, immunostaining, microscopy and image analysis were performed as described previously (Amin et al., 2007). Guinea pig anti-FOZI-1 (1:200) and goat anti-GFP (Rockland Immunochemicals; 1:1000) were used. All secondary antibodies were from Jackson ImmunoResearch Laboratories and used at a dilution of 1:50 to 1:200.

3.2.8 Fractionation experiment

Transgenic worms used in the fractionation experiment were generated in the *pha-1(e2123ts)* mutant background and maintained at 25°C. Worms were harvested in M9 and re-suspended in lysis buffer [50mM Tris-HCI, pH=7.5; 10% glycerol; 10mM 2-mercaptoethanol, 1mM phenylmethylsulfonyl fluoride and 1 Complete Mini Protease Inhibitor Cocktail Tablet (Roche)/50ml lysis buffer] so that the ratio of worm volume to buffer volume was 1:1. Worms were sonicated until around 90% of the worms were lysed as assessed by microscopy. Debris was cleared by centrifugation at 4000xg for 10min at 4°C. The supernatants were centrifuged at 115,000xg for 1h. The resulting supernatant was the soluble

fraction. The pellet, which contains microsomal membrane proteins, was further cleared from soluble proteins by two more rounds of centrifugation (115,000xg)/resuspension in lysis buffer. The resulting pellet was reconstituted in the same volume of lysis buffer as the soluble fraction. Micro BCA Protein Assay Reagent Kit (Pierce) was used to measure the protein concentration. 30mg/lane of samples were loaded onto a 7% SDS-PAGE. GFP fusion proteins were detected by Western blotting using goat anti-GFP antibodies (Rockland Immunochemicals; 1:1000).

3.2.9 Generating the RAD-SMAD and BAD-SMAD reporters:

Oligonucleotides containing the wild-type (RAD-SMAD) or mutant (BAD-SMAD) Smad binding sites and StyI restriction sites at each end were treated with T4 polynucleotide kinase, annealed and ligated. Concatenated oligonucleotides were cloned into the Fire lab vector L3135 (http://www.addgene.org). Upon conformation of plasmid sequence, their DNA was injected into N2 worms using *mec-7p::mRFP* as a co-injection marker. The RAD-SMAD transgene was then integrated into the genome via gamma-irradiation. Seven integrated lines were obtained, out-crossed and mapped to specific chromosomes. All lines showed similar GFP expression patterns. Two of them, LW2433: *jjIs2433[pCXT51(RAD-SMAD) + LiuFD61(mec-7p::mRFP)]* X and LW2436: *jjIs2436[pCXT51(RAD-SMAD) + LiuFD61(mec-7p:: mRFP)]* (I or IV) were used for subsequent analysis. The BAD-SMAD transgenes were kept in high efficiency extra chromosomal lines.

3.3 RESULTS

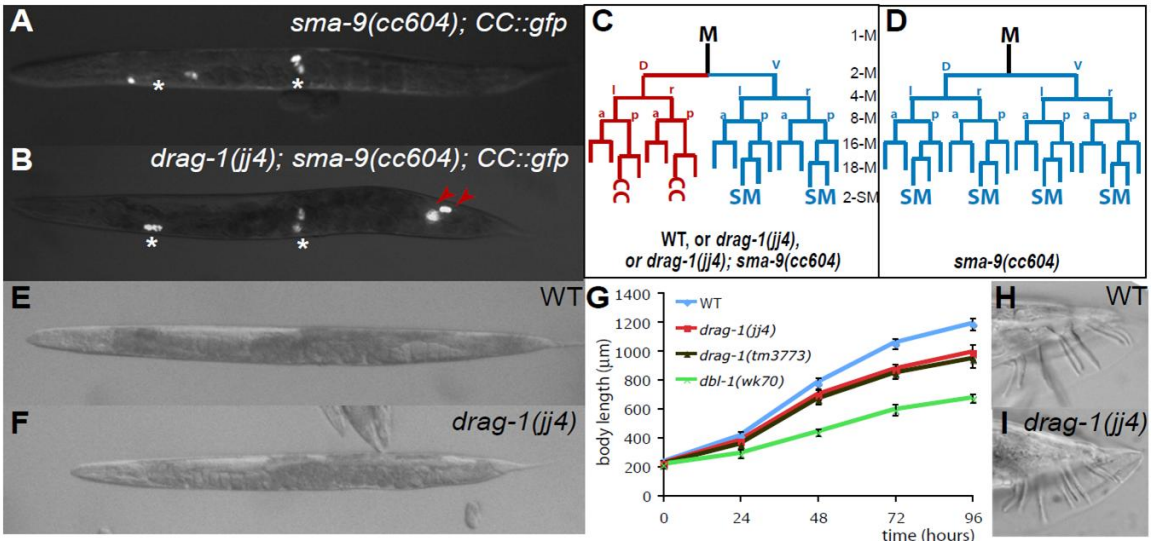
3.3.1 drag-1 mutants exhibit a subset of the phenotypes seen in mutants in the Sma/Mab pathway

In a mutagenesis screen for suppressers of the *sma-9(cc604)* M lineage phenotype (Foehr et al., 2006), we isolated the Sma/Mab pathway mutants as well as a single locus, recessive mutation *jj4*. Unlike *sma-9(cc604)* animals, which lack both M-lineage derived CCs (Figure 3.1A,D), 98.5% (n=201) of *jj4; sma-9(cc604)* animals had both M-derived CCs (Figure 3.1B,C). *jj4* also suppressed the M lineage phenotypes of two other *sma-9* alleles, *cc606* and *wk5* (data not shown), suggesting that this suppression is not allele specific. Like the Sma/Mab pathway mutants, *jj4* mutants by themselves did not have any M lineage defect when separated from the *cc604* mutation (Figure 3.1C). However, *jj4* mutants were smaller than wild-type N2 worms throughout larval development, although they were not as small as the *dbl-1(wk70)* null mutants (Figure 3.1E-G). Unlike the Sma/Mab pathway mutants, *jj4* mutant males could mate and did not exhibit any male tail patterning defects (Figure 3.1H,I). Thus, *jj4* mutants exhibit some, but not all, phenotypes of the Sma/Mab pathway mutants.

We mapped the *jj4* mutation to the left arm of chromosome I, a region not previously known to contain any genes involved in body size regulation. Thus *jj4* represents a novel locus required to regulate body size. We named the locus *drag-1* (see below).

Figure 3.1: *drag-1* mutants exhibit body size and mesodermal defects.

All images are ventral/lateral views with anterior to the left. *CC::gfp* labels all the coelomocytes (CCs) , including embryonically-derived CCs (*) and M-derived CCs (red arrowheads). (A-D) Mesoderm phenotype. *sma-9(cc604)* mutants (A, D) show a dorsal to ventral fate transformation in the postembryonic mesoderm lineage, the M lineage, thus lacking M-derived CCs, compared to wild-type (C) or *drag-1(jj4); sma-9(cc604)* double mutants (B, C). (E-F) A wild-type adult (E) and *drag-1(jj4)* adult (F) at the same stage. (G) Growth curves of wild-type (N2), *drag-1(jj4)*, *drag-1(tm3773)* and *dbl-1(wk70)* worms. Fifty to 85 animals were examined for each strain at each time point. By student t test, the body lengths of *drag-1(jj4)* and *drag-1(tm3773)* animals are significantly different (p<0.0001) from the body lengths of wild-type and *dbl-1(wk70)* animals at 24, 48, 72 and 96h. (H, I) Tails of a wild-type male (H) and a *drag-1(jj4)* male (I). SM, sex myoblast. a, anterior; d, dorsal; l, left; p, posterior; r, right; v, ventral.


A
sma-9(cc604); CC::gfp
B
drag-1(jj4); sma-9(cc604); CC::gfp
C
M
D
V
1-M
2-M
4-M
8-M
16-M
18-M
2-SM
SM
WT, or drag-1(jj4),
or drag-1(jj4); sma-9(cc604)
D
sma-9(cc604)
E
WT
F
drag-1(jj4)
G
WT
drag-1(jj4)
drag-1(tm3773)
dbl-1(wk70)
body length (µm)
time (hours)
H
WT
I
drag-1(jj4)

3.3.2 drag-1 functions in the Sma/Mab pathway, possibly at the ligand-receptor level, to regulate body size

Because *drag-1(jj4)* mutants and the Sma/Mab pathway mutants share similar phenotypes in body size and mesoderm patterning, we tested whether *drag-1* functions in the Sma/Mab pathway. We generated double mutants between *jj4* (a putative null, see below) and mutations in various Sma/Mab pathway components (Figure 3.2A) and measured their body sizes.

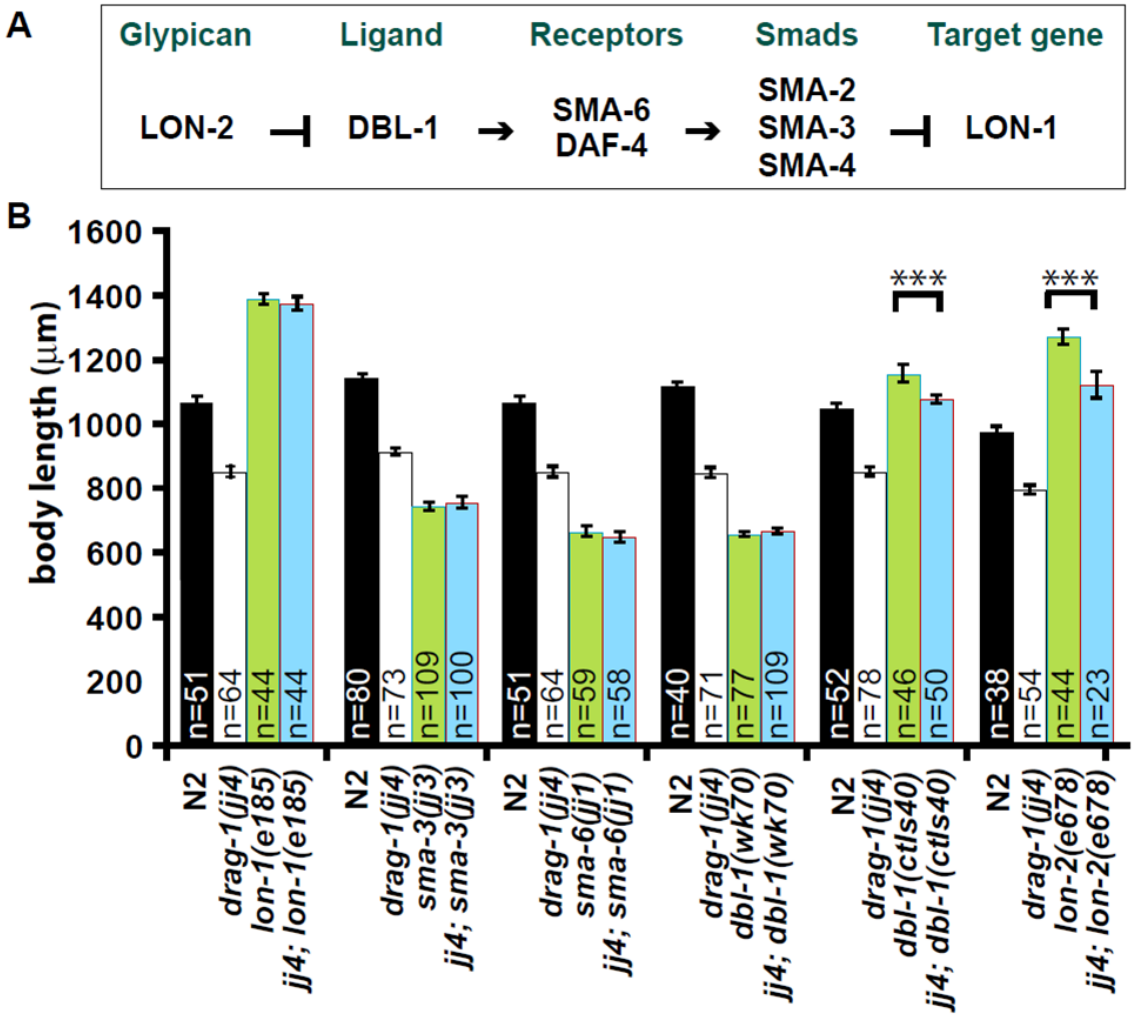
As shown in Figure 3.2B, double mutants between *jj4* and the *sma-3(jj3)* null mutation were as small as *sma-3(jj3)* single mutants. Similarly, *jj4; sma-6(jj1)* and *jj4; dbl-1(wk70)* double mutants were as small as *sma-6(jj1)* and *dbl-1(wk70)* single mutants (both are null mutants), respectively. These observations suggest that *jj4* is likely compromised for Sma/Mab pathway function, rather than affecting a pathway that functions in parallel to the Sma/Mab pathway, in regulating body size.

To further characterize the role of *drag-1* in the Sma/Mab pathway, we generated the following double mutants: *jj4; lon-1(e185)*, *jj4; lon-2(e678)* and *jj4; dbl-1(ctIs40)*. *e185* is a strong loss-of-function mutation in *lon-1*, a known downstream target of the Sma/Mab pathway (Maduzia et al., 2002; Morita et al., 2002). *e678* is a null mutation in *lon-2*, which encodes a member of the glypican family of heparan sulfate proteoglycans and acts as a negative regulator of the Sma/Mab pathway (Gumienny et al., 2007). *dbl-1(ctIs40)* is a strain that over-expresses *dbl-1* (Suzuki et al., 1999). As shown in Figure 3.2B, *jj4; lon-1(e185)*

mutants were as long as *lon-1(e185)* mutants. However, *jj4; lon-2(e678)* and *jj4; dbl-1(ctIs40)* double mutants showed intermediate body size between *jj4* single mutants and *lon-2(e678)* or *dbl-1(ctIs40)* single mutants. These results suggest that *drag-1* is likely to act upstream of *lon-1*, but in parallel to *lon-2* and *dbl-1*. Taken together, our genetic epistasis results are consistent with *drag-1* functioning at the ligand/receptor level in the Sma/Mab pathway to regulate body size.

Figure 3.2: *drag-1* likely functions at the ligand-receptor level in the Sma/Mab pathway to regulate body size.

(A) A simplified schematic of the Sma/Mab pathway (see text for details). (B) Body length of wild-type (N2) and various mutant worms at 96hr post plating (see Materials and Methods). Similar results were observed using worms at 24hr, 48hr and 72hr post plating (data not shown). *drag-1(jj4); lon-2(e678)* and *drag-1(jj4); dbl-1(ctIs40)* worms show significantly different body length compared to *lon-2(e678)* and *dbl-1(ctIs40)* worms, respectively. Error bars represent 95% confidence intervals for the mean body length. The significance of difference between double mutant and the corresponding single mutant was statistically analyzed via student *t* test. ***, $p<0.0001$.


A
Glypican
Ligand
Receptors
Smads
Target gene
LON-2
DBL-1
SMA-6
DAF-4
SMA-2
SMA-3
SMA-4
LON-1
B
body length (μm)
1600
1400
1200
1000
800
600
400
200
0
***
***
n=51
n=64
n=44
n=44
n=80
n=73
n=109
n=100
n=51
n=64
n=59
n=58
n=40
n=71
n=77
n=109
n=52
n=78
n=46
n=50
n=38
n=54
n=44
n=23
N2
drag-1(jj4)
lon-1(e185)
jj4; lon-1(e185)
N2
drag-1(jj4)
sma-3(jj3)
jj4; sma-3(jj3)
N2
drag-1(jj4)
sma-6(jj1)
jj4; sma-6(jj1)
N2
drag-1(jj4)
dbl-1(wk70)
jj4; dbl-1(wk70)
N2
drag-1(jj4)
dbl-1(ctls40)
jj4; dbl-1(ctls40)
N2
drag-1(jj4)
lon-2(e678)
jj4; lon-2(e678)

3.3.3 drag-1 interacts genetically with the dauer pathway

In addition to the Sma/Mab pathway, *C. elegans* has another TGFβ pathway that regulates dauer development (Savage-Dunn, 2005). *drag-1(jj4)* animals did not exhibit any constitutive or deficient dauer phenotypes on their own (Table 3.1). Animals carrying a deletion allele of *drag-1*, *tm3773* (see below), did not exhibit any dauer phenotypes either (Table 3.1). However, *drag-1(jj4)* enhanced the dauer-constitutive (Daf-c) phenotype of certain *daf-7* and *daf-1* alleles at the permissive temperature. *daf-7* and *daf-1* encode the ligand and the type I receptor of the dauer pathway, respectively (Georgi et al., 1990; Ren et al., 1996). As shown in Table 3.1, while *drag-1(jj4)* did not affect the Daf-c phenotype of *daf-7(m62)* animals at all temperatures tested, *drag-1(jj4)* moderately enhanced the Daf-c phenotype of *daf-7(e1372)* mutants at 16˚C. Similarly, *drag-1(jj4)* moderately enhanced the Daf-c phenotype of *daf-1(m213)* mutants at 16˚C, but significantly enhanced the Daf-c phenotype of *daf-1(m40)* mutants at both 16˚C and 20˚C. The allele specific and moderate enhancement of Daf-c phenptypes of *daf-7* and *daf-1* mutants by *drag-1(jj4)* are consistent with low levels of crosstalk between the Sma/Mab and the dauer pathways, as suggested by previous observations on the genetic interactions between *sma-6* and *daf-7* and *daf-1* mutations (Krishna et al., 1999).

Table 3.1: Enhancement of Daf-c phenotypes of *daf-1* and *daf-7* mutants by the *drag-1* mutation

	15˚C		20˚C		25˚C	
	n	% Daf-c (pop)	n	% Daf-c (pop)	n	% Daf-c (pop)
drag-1(jj4)	3	0 (476)	3	0 (395)	3	0 (275)
drag-1(tm3773)	3	0 (167)	3	0 (381)	3	0 (209)
daf-7(m62)	5	79.5±2.9 (317)	6	80.0±12.2	5	100 (373)
drag-1(jj4); daf-7(m62)	5	80.0±3.5 (177)	6	77.1±3.7 (357)	5	100 (502)
daf-7(e1372)	5	63.3±2.7 (296)	5	73.3±10.7	5	100 (613)
drag-1(jj4); daf-7(e1372)	5	68.7±3.5	5	81.5±6.3 (636)	5	100 (627)
daf-1(m40)	4	3.1±1.3 (378)	6	20.5±7.3 (991)	5	100 (570)
drag-1(jj4); daf-1(m40)	5	27.9±9.6	5	95.1±3.4	5	100 (470)
daf-1(m213)	5	0 (302)	5	80.9±6.4 (606)	5	100 (780)
drag-1(jj4); daf-1(n213)	5	2.1±1.9 (459)*	5	81.5±8.9 (255)	5	100 (1096)

n: the number of plates scored at each temperature;

% Daf-c: the mean and standard deviation of the dauer formation percentage for n plates;

pop: total number of worms on n plates scored;

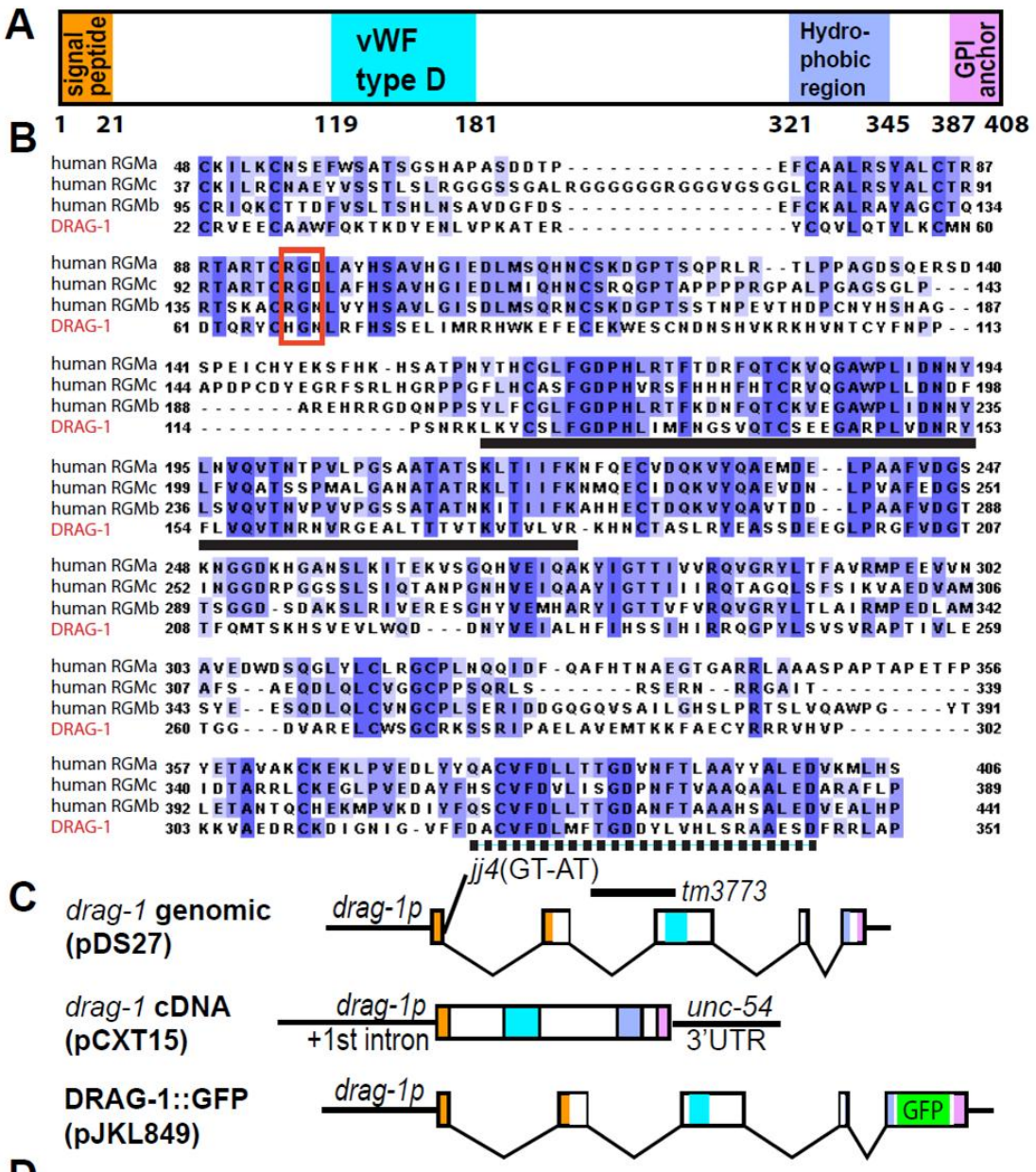
*: $p<0.05$, as calculated by t test, between double mutant and the corresponding *daf* single mutant at the specified temperature.

3.3.4 drag-1 encodes the C. elegans homolog of the RGM family proteins

drag-1(jj4) was mapped to the left arm of chromosome I between Y92H12A.3 and Y71G12B.24 via snip-SNP mapping (Wicks et al., 2001). RNAi injection for 15 likely candidates among the 24 predicted genes located in this region showed that *Y71G12B.16(RNAi)* led to a small body size phenotype and the suppression of the *sma-9(cc604)* M lineage phenotype. For simplicity, we will refer to the suppression of the *sma-9(cc604)* M lineage phenotype by *jj4* as the SOSMLP phenotype of *jj4* in the remainder of the manuscript. Several additional lines of evidence indicated that *jj4* is a loss-of-function allele of Y71G12B.16. First, DNA fragments containing the genomic DNA sequence of Y71G12B.16 rescued both the small body size and the SOSMLP phenotype of *jj4* mutants (Figures 3.3C,D and 3.5O). Second, sequencing of Y71G12B.16 in *jj4* mutants revealed a T to A single nucleotide mutation in the splicing donor site in the first intron (Figure 3.3C). RT-PCR analysis (see Materials and Methods) showed that this mutation led to the production of three aberrant *drag-1* transcripts, all producing mutant peptides containing the first 16 amino acids of DRAG-1 followed by 9-49 additional random amino acids (data not shown). Third, a deletion allele of *drag-1*, *tm3773,* led to the same body size and SOSMLP phenotypes as *jj4* mutants (Figures 3.1G, 3.3D and 3.5O).

Figure 3.3: *drag-1* encodes a putative GPI-anchor protein of the RGM family.

(A) A schematic of the DRAG-1 protein, showing the various conserved motifs drawn in scale. (B) Alignment between DRAG-1 (the bottom line) and its human RGM homologs RGMa (GenBank: AAI51133.1), RGMb (GenBank: NP_001012779.2) and RGMc (GenBank: NP_998818.1) in regions between the N-terminal signal peptide and the C-terminal pro-peptide. Notice the highly conserved vWF type D domain (solid underline) and hydrophobic region (dashed underline). The RGD motif (boxed in red) is present in RGMa and RGMc, but not in RGMb and DRAG-1. (C) Diagrams of the *drag-1* genomic, cDNA and GFP tagged constructs. The locations of the *jj4* and *tm3773* molecular lesions are shown. (D) Body size measurement of worms of various genotypes at 72hr post-plating. Length is presented as the mean +/- standard deviation. * indicates the length of the *drag-1* worms carrying a particular transgene is significantly different from that of the corresponding *drag-1* single mutant. $p<0.0001$.



D

Genotype	Length (mm)	N
wild-type (N2)	0.91 +/- 0.06	40
drag-1(jj4)	0.81 +/- 0.05	83
drag-1(jj4); Ex[pDS27(drag-1 genomic)]	0.93 +/- 0.05*	75
drag-1(jj4); Ex[pCXT15(drag-1 cDNA)]	0.90 +/- 0.05*	57
drag-1(jj4); Ex[pJKL849(DRAG-1::GFP)]	0.93 +/- 0.06*	60
drag-1(tm3773)	0.81 +/- 0.05	36
drag-1(tm3773); Ex[pDS27(drag-1 genomic)]	0.97 +/- 0.05*	25

We isolated two cDNA variants for Y71G12B.16, Y71G12B.16a and Y71G12B.16b, which share the same coding sequences and differ only in the lengths of their 5' UTRs (see Materials and Methods). The predicted Y71G12B.16 protein is 408aa long. It contains a predicted N-terminal signal peptide, a partial von Willebrand factor D (vWF type D) domain, a hydrophobic region and a C-terminal region that meets the criteria of a pro-peptide that is cleaved off and replaced by a glycosylphosphatidylinositol (GPI) anchor during post-translational processing (Chatterjee and Mayor, 2001; Bohme and Cross, 2002). Blast search showed that Y71G12B.16 has homologs in sea urchins, tunicates, mollusks and vertebrates, but not in Drosophila. The vertebrate Y71G12B.16 homologs belong to the RGM (Repulsive Guidance Molecule) family that includes RGMa, RGMb (aka DRAGON), RGMc (aka hemojuvelin) and the recently identified RGMd (Camus and Lambert, 2007; Corradini et al., 2009). Y71G12B.16 appears to be a distant member of the RGM family (Camus and Lambert, 2007), sharing 22%, 21% and 18% identity to Human RGMb, RGMa and RGMc respectively, in the region flanked by the N-terminal signal peptide and the C-terminal pro-peptide (Figure 3.3B). In particular, Y71G12B.16 and human RGMb/DRAGON share 46% identity in the vWF type D domain and 40% identity in the hydrophobic region. Like RGMb/DRAGON, Y71G12B.16 lacks the RGD (Arg-Gly-Asp) motif found in RGMa and RGMb proteins (Figure 3.3B). We therefore named Y71G12B.16 DRAG-1. Like *C. elegans*, other invertebrates including sea urchins, tunicates and mollusks also contain a single RGM-related protein in each genome (Camus and Lambert, 2007).

RGM proteins can function as BMP co-receptors (Babitt et al., 2005, 2006; Samad et al., 2005). Our genetic evidence described above suggests that DRAG-1 may also function as a co-receptor in the Sma/Mab pathway in *C. elegans*.

3.3.5 drag-1 is expressed in the same cell types as the Sma/Mab pathway type I receptor sma-6

To test the hypothesis that DRAG-1 may function as a co-receptor for the Sma/Mab pathway, we first asked whether *drag-1* is expressed in the same cells that express core components of the Sma/Mab pathway. A *drag-1* genomic construct (pDS27), which contains the entire *drag-1* genomic region including 4kb 5' sequences and 1kb 3' sequences, fully rescued the body size defects of both *drag-1(jj4)* and *drag-1(tm3773)* mutants (Figure 3.3C,D). Similarly, a *drag-1* cDNA construct (pCXT15) (see Materials and Methods), with the genomic coding region of *drag-1* replaced by *drag-1* cDNA and the *drag-1* 3' UTR replaced by *unc-54* 3' UTR, also rescued the body size defects of *drag-1(jj4)* mutants (Figure 3.3C,D). These results suggest that the regulatory elements required for *drag-1* function in regulating body size likely reside in the upstream sequences used in pCXT15. We then generated a transcriptional *drag-1p::gfp* fusion pDS15 using these upstream sequences (see Materials and Methods). Transgenic lines carrying pDS15 showed GFP expression in pharyngeal, hypodermal and intestinal cells (Figure 3.4A-C), the same cells that express *sma-6* in hermaphrodites (Krishna et al., 1999; Yoshida et al., 2001). In fact, transgenic

lines carrying both *drag-1p::gfp* and *sma-6p::mRFP* showed co-localization of the two reporters in all three tissues described above (Figure 3.4A-C, E-G, I-K). Therefore, *drag-1* is expressed in the same cell types as *sma-6* in hermaphrodites. However, *drag-1p::gfp* is not expressed in the male tail cells that express *sma-6p::mRFP* (Figure 3.4D, H and L), consistent with the lack of male tail defects in *drag-1(jj4)* mutant males.

The same expression pattern of *drag-1* was also observed using a functional translational DRAG-1::GFP fusion (pJKL849), which has GFP inserted right before the putative cleavage site upstream of the C terminal pro-peptide sequence (Figure 3.3B). This reporter is fully functional as it rescued both the body size and the SOSMLP of *jj4* mutants (Figures 3.3D, 3.5D,O). Just like the *drag-1p::gfp* transcriptional fusion, DRAG-1::GFP was present in pharyngeal, hypodermal and intestinal cells (Figure 3.4M-O). Robust expression in the M lineage was also detected from 1-M to 4-M stage (Figure 3.4Q-S). DRAG-1::GFP signal was reduced (40%, n=20) or undetectable (60%, n=20) at the 8-M stage (Figure 3.4T), and completely undetectable beyond the 8-M stage (Figure 3.4U and data not shown). The M lineage expression pattern of *drag-1* is consistent with previous findings that *sma-9* functions prior to the 8-M stage for proper M lineage development (Foehr et al., 2006).

3.3.6 drag-1 functions in the same cells as the Sma/Mab pathway receptors and Smads to regulate body size and mesoderm patterning

The Sma/Mab pathway receptors and Smad proteins have been shown to function in the hypodermal cells to regulate body size (Yoshida et al., 2001; Wang et al., 2002). Because *drag-1* is expressed in the same tissues as *sma-6* in hermaphrodites, we tested whether *drag-1* also functions in hypodermal cells to regulate body size. We forced the expression of *drag-1* cDNA in pharyngeal muscles [using the *myo-2* promoter (Okkema et al., 1993)], intestine [using the *elt-2* promoter (Fukushige et al., 1998)], and in hypodermal cells [using both the *elt-3* and *rol-6* promoters (Kramer and Johnson, 1993; Gilleard et al., 1999)], and tested the ability of each transgene to rescue the small body size of *jj4* mutants. As a control, *drag-1* cDNA under the control of its own promoter could rescue the small body size phenotype of *jj4* mutants (Figure 3.4P). As shown in Figure 3.4P, forced expression of *drag-1* in hypodermal cells rescued the small body size phenotype of *jj4* mutants, while expression in the pharynx and intestine did not. Thus, *drag-1* functions in the same cells as the Sma/Mab receptors and Smad proteins to regulate body size.

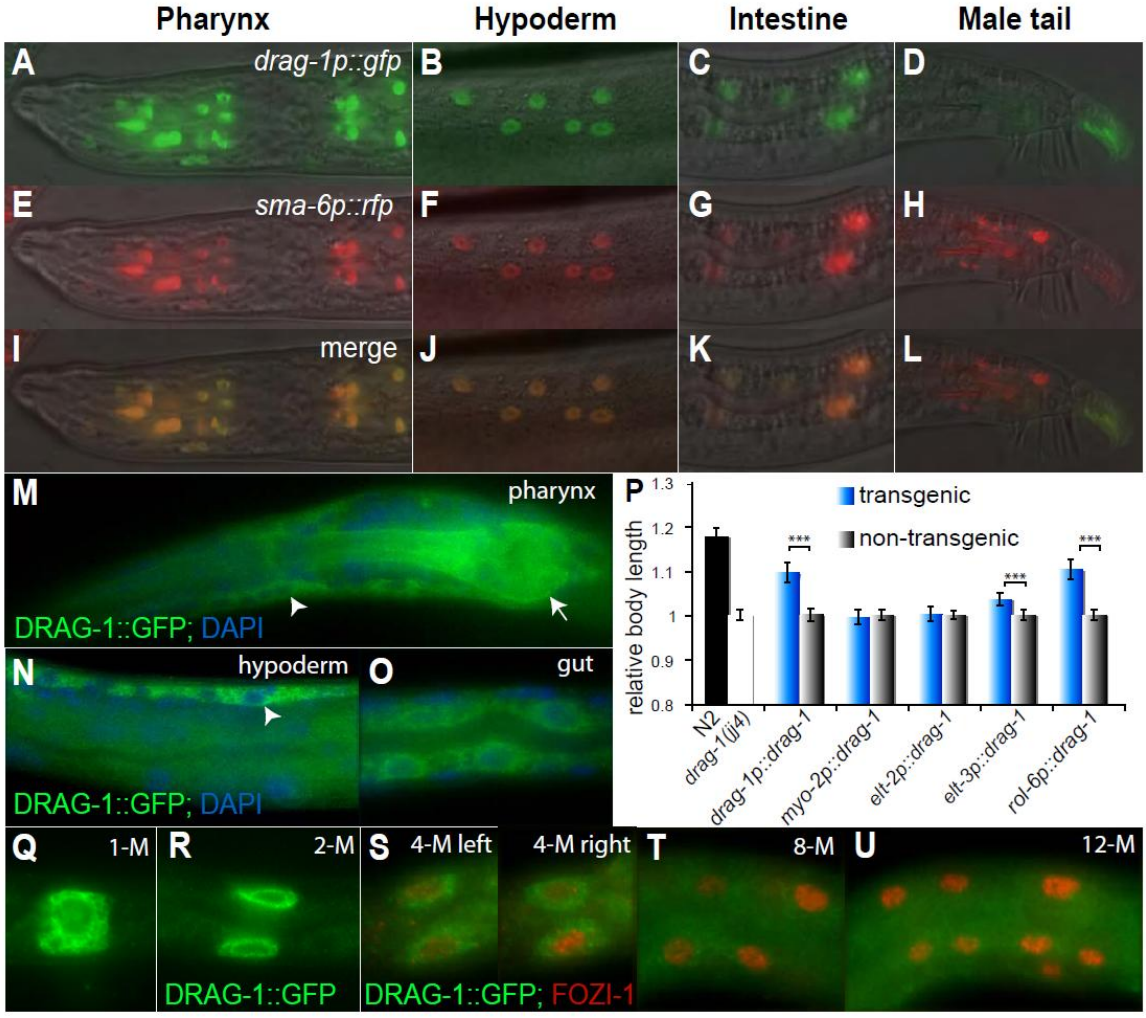
Previous studies have shown that *sma-9* and the Sma/Mab pathway function within the M lineage to regulate dorsoventral patterning of the M lineage (Foehr et al., 2006). We therefore forced the expression of *drag-1* within the M lineage in *drag-1(jj4); sma-9(cc604)* double mutants using the *hlh-8* promoter (Harfe et al., 1998). This M lineage-specific expression of *drag-1* (using pCXT148) is sufficient to rescue the SOSMLP phenotype of *jj4* mutants (24.5%, n=106), i.e.,

to reverse the M lineage phenotype of *drag-1(jj4); sma-9(cc604)* mutants to that of *sma-9(cc604)*. Thus, *drag-1* also functions in the same cells as *sma-9* and the Sma/Mab pathway to regulate M lineage development.

Figure 3.4: *drag-1* is expressed and functions in the same cell types as the Sma/Mab pathway components.

Images shown are side views with anterior to the left and posterior to the right. *drag-1p::gfp* (A-D) and *sma-6p::rfp* (E-H) co-localize in pharyngeal (A, E, I), hypodermal (B, F, J) and intestinal (C, G, K) cells, but do not co-localize in the male tail (D, H, L). (I-L) are merged images of (A-D) and (E-H) respectively. (M-N) DRAG-1::GFP or (O) LIN-12TM::GFP (see Figure 3.5 for details) localization in pharyngeal (M), hypodermal (N) and intestinal (O) cells, as visualized by anti-GFP antibody staining. Note that the GFP signal (green) is located outside of the nucleus (blue, stained with DAPI), both at the cell surface and inside the cell. Arrow heads, hypodermal cells; arrow, pharynx. (P) Tissue-specific rescue of the body size phenotype of *drag-1(jj4)* mutants. Body size was measured in adult worms at 72hr post-plating. Transgenic worms were distinguished from non-transgenic ones by the presence of *mec-7p::rfp*. The lengths of *drag-1(jj4)* and non-transgenic worms of each group were normalized to 1. The relative lengths of wild-type and various transgenic animals compared to their corresponding non-transgenic controls are presented. Between 30 and 80 animals were measured for each genotype. Error bars represent 95% confidence intervals for the mean of relative body length. The significance of difference between transgenic and the corresponding non-transgenic animals was statistically analyzed via student *t* test. ***, $p<0.0001$. (Q-U) The M lineage expression pattern of *drag-1* using LIN-12TM::GFP (green, see Figure 3.5 for details). Anti-

FOZI-1 antibody staining (red) was used to mark M lineage cells from the 4-M to the 12-M stage (S-U). Only one focal plane was shown for the 8-M (T) and 12-M (U) stage worms. DRAG-1::GFP is present from the 1-M to the 4-M stage (O-S), becomes fainter at the 8-M stage (T) and undetectable after the 8-M stage (U).


Pharynx
Hypoderm
Intestine
Male tail
A
drag-1p::gfp
B
C
D
E
sma-6p::rfp
F
G
H
I
merge
J
K
L
M
pharynx
DRAG-1::GFP; DAPI
N
hypoderm
O
gut
DRAG-1::GFP; DAPI
P
transgenic
non-transgenic
relative body length
1.3
1.2
1.1
1
0.9
0.8
***
***
***
N2
drag-1(ji4)
drag-1p::drag-1
myo-2p::drag-1
elt-2p::drag-1
elt-3p::drag-1
rol-6p::drag-1
Q
1-M
R
2-M
DRAG-1::GFP
S
4-M left
4-M right
DRAG-1::GFP; FOZI-1
T
8-M
U
12-M

3.3.7 DRAG-1 localizes to and functions at the cell membrane

The translational DRAG-1::GFP reporter described above allowed us to examine the sub-cellular localization pattern of DRAG-1. As shown in Figures 3.4M-O, 3.5B-C, DRAG-1::GFP was present outside of the nucleus but localized to the cell surface as well as inside the cell. To determine if DRAG-1 protein is membrane-associated, we performed cell fractionation experiments to separate the soluble fraction and the membrane fraction using lysates from mix-staged worm populations (see Materials and Methods). As controls, we used worms expressing GFP alone under the control of the *hmt-1* promoter, and worms expressing the HMT-1::GFP fusion under the control of the *hmt-1* promoter (Schwartz et al., 2010; see Materials and Methods). HMT-1 is a transmembrane half-molecule ATP-binding cassette transporter required for heavy metal detoxification (Vatamaniuk et al., 2005). As shown in Figure 3.5N, GFP was only detected in the soluble fraction, while HMT-1::GFP was only detected in the membrane fraction. In the same experiment, DRAG-1::GFP was only detected in the membrane fraction. This result, in combination with results from immunostaining, demonstrates that DRAG-1 is a membrane protein, as predicted for a putative GPI-anchor protein.

To determine whether DRAG-1 functions at the cell membrane, we generated transgenic lines expressing DelC::GFP, which has the C-terminal pro-peptide deleted from DRAG-1 (Figure 3.5A). DelC::GFP showed reduced signal at the cell surface (Figure 3.5E,F), and became partially soluble in cell fractionation experiments (Figure 3.5N). We then checked the functionality of

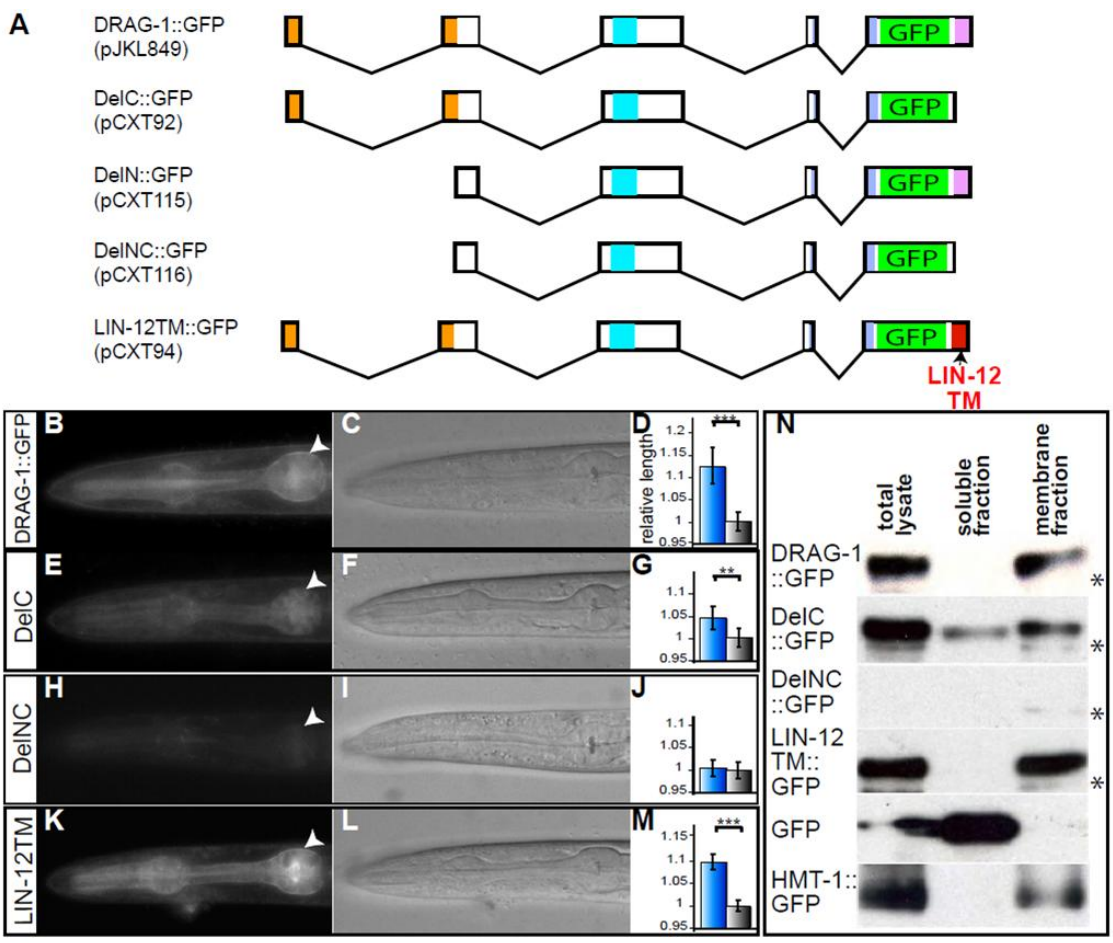
DelC::GFP and found that it could rescue both the body size and the SOSMLP phenotypes of *drag-1(jj4)* mutants, but the rescuing efficiency was reduced when compared to wild-type DRAG-1::GFP (Figures 3.5D,G,O). These data together suggest that the putative C-terminal pro-peptide is important, but not essential, for DRAG-1 membrane localization and function.

The membrane localization and function of DRAG-1::GFP and DelC::GFP appeared to require that the proteins enter the secretory pathway via their N-terminal signal peptide. DelN::GFP and DelNC::GFP with their N-terminal signal peptide deleted (Figure 3.5A) failed to rescue the *drag-1(jj4)* mutant phenotypes (Figure 3.5J, O), and showed drastically reduced level of expression (Figure 3.5H,I,N), possibly due to the mis-targeting and subsequent degradation, of the mutant protein.

To directly test whether DRAG-1 functions at the cell membrane, we replaced the putative C-terminal pro-peptide sequence of DRAG-1 with the transmembrane domain of a well-characterized transmembrane protein LIN-12, LIN-12TM::GFP (Figure 3.5A). LIN-12TM::GFP was solely detected in the membrane fraction (Figure 3.5N), showed cell surface localization (Figure 3.5K,L), and fully rescued the body size and SOSMLP phenotypes of *drag-1(jj4)* mutant (Figure 3.5M,O). Thus DRAG-1 is not only present, but also functions, at the cell membrane.

Figure 3.5: *drag-1* is localized to and functions at the cell membrane.

(A) Schematics of various DRAG-1 deletion constructs (see Material and Methods). (B-M) GFP localization (B, E, H, K, arrowheads point to the surface of the pharynx), the corresponding DIC images (C, F, I, L), and body size measurement (D, G, J, M) of animals containing the specific transgenes. (B, C, D) DRAG-1::GFP, (E, F, G) DelC::GFP (abbreviated as DelC), (H, I, J) DelNC::GFP (abbreviated as DelNC), and (K, L, M) LIN-12TM::GFP (abbreviated as LIN-12TM). In panels D, G, J and M, Blue bars: *drag-1(jj4)* animals carrying the transgene; grey bars: *drag-1(jj4)* non-transgenic animals. the Y axis shows the relative body length, with the lengths of non-transgenic worms of each group being normalized to 1. Between 30 and 70 animals were measured for each genotype. Error bars represent 95% confidence intervals for the mean of relative body length. The significance of difference between transgenic and the corresponding non-transgenic animals was statistically analyzed via student *t* test. ***, $p<0.0001$. **, $p<0.001$. (N) Western blots probed with anti-GFP antibodies showing the localization of DRAG-1::GFP fusions via fractionation experiments. * refers to non-specific bands recognized faintly by the anti-GFP antibodies. (O) Rescue of the mesodermal phenotype of *jj4* mutants by various *drag-1::gfp* constructs.



O

Genotype	0 M-CC (%)	2 M-CC (%)	N
wild type	0	100	500
sma-9(cc604)	97	3	103
drag-1(jj4); sma-9(cc604)	1.5	98.5	201
drag-1(tm3773); sma-9(cc604)	1.0	99.0	400
drag-1(jj4); sma-9(cc604); Ex[pDS27(DRAG-1 genomic)]	94.2	5.8	225
drag-1(jj4); sma-9(cc604); Ex[pJKL849(DRAG-1::GFP)]	100	0	43
drag-1(jj4); sma-9(cc604); Ex[pCXT92(DelC)]	55.5	44.5	164
drag-1(jj4); sma-9(cc604); Ex[pCXT115(DelN)]	5.4	94.6	129
drag-1(jj4); sma-9(cc604); Ex[pCXT116(DelNC)]	3.6	96.4	56
drag-1(jj4); sma-9(cc604); Ex[pCXT94(LIN-12TM)]	79.4	20.6	214

3.3.8 DRAG-1 positively modulates Sma/Mab signaling as indicated by a Sma/Mab-responsive reporter

The phenotypes of *drag-1* mutants described above suggest that DRAG-1 positively modulates Sma/Mab signaling. To test this hypothesis, we generated a Sma/Mab-responsive reporter as there are no existing phospho-Smad antibodies or appropriate Smad::GFP reporters or anti-Smad antibodies that allow us to directly monitor the output of Sma/Mab signaling.

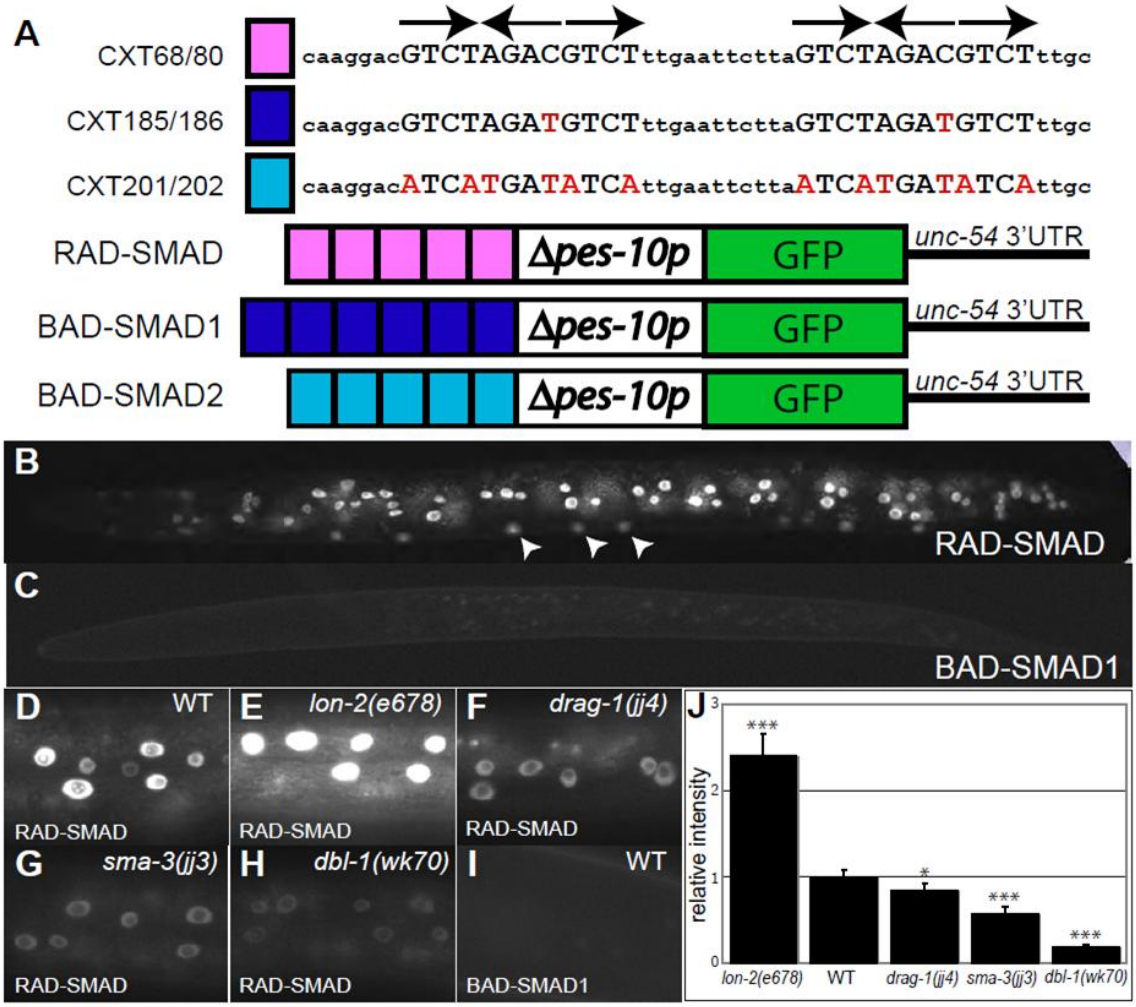
Previous work has shown that individual Smad complexes are capable of binding to three abutting Smad boxes (or Smad binding site: GTCT) with high affinity (Johnson et al., 1999). Furthermore, abutting Smad boxes organized in the RLR orientation (R, rightward and L, leftward) are bound by Smad complexes with higher affinity than those organized in the RLL or RRR orientation (Johnson et al., 1999). We decided to place multiple copies of the Smad boxes organized in the RLR orientation upstream of the minimal *pes-10* promoter and *gfp* (Figure 3.6A), and tested whether the expression of the *gfp* reporter is responsive to Sma/Mab signaling. We named the reporter as RAD-SMAD (Reporter acting downstream of Smad). The expression pattern of the *gfp* reporter was examined using two different integrated transgenic lines carrying RAD-SMAD (see Materials and Methods), which showed similar expression patterns. We detected GFP signal from late embryogenesis until adulthood. GFP signal was detected in intestinal and hypodermal cells, including those from the P lineages (Figure 3.6B) and cells in the male tail (data not shown). The expression level of the reporter appeared rather dynamic during development, with stronger intestinal expression

in embryos and L1s but fainter in adults, and stronger hypodermal expression in L2 and L3 stages (data not shown). No *gfp* expression was detected in dauer animals either in the wild-type background or the *daf-7(e1372)* background (data not shown).

To test the specificity of the RAD-SMAD reporter, we generated two reporters carrying different mutations in the Smad boxes, BAD-SMAD1 and BAD-SMAD2 (Bad reporter of Smad) (Figure 3.6A). Transgenic worms carrying either reporter showed no *gfp* expression in any cells (Figure 3.6C, I, and data not shown). This result indicates that the Smad boxes are directly responsible for RAD-SMAD reporter expression.

Figure 3.6: A RAD-SMAD reporter directly and positively responds to Sma/Mab signaling *in vivo*.

(A) Schematics of the RAD-SMAD, BAD-SMAD1 and BAD-SMAD2 reporters, with the oligonucleotide sequences shown on top. Mutated sequences in the BAD-SMAD1 and BAD-SMAD2 reporters are shown in red lower case. (B, D-H) Expression of the RAD-SMAD GFP reporter in wild-type (B, D), *lon-2(e678)* (E), *drag-1(jj4)* (F), *sma-3(jj3)* (G) and *dbl-1(wk70)* (H) worms at the L2 stage. Arrowheads in B point to hypodermal cells derived from the P lineages. (C, I) The BAD-SMAD1 GFP reporter shows no expression. Hypodermal cells are shown in (D-I). Pictures were taken at the same exposure. (J) Quantification of hypodermal RAD-SMAD GFP signals in various mutants. In each mutant background, the pixel intensities of GFP signals from 10 hypodermal nuclei in 10 different worms (20 worms for WT) at the L2 stage were measured by the OpenLab software. The signal intensity in wild-type worms was set to 1 and compared to the signal intensity in each mutant background. Error bars represent 95% confidence intervals for the means of relative intensity. The significance of difference between wild-type and each mutant was statistically analyzed via student *t* test. *** $p<0.0001$; * $p<0.01$.


A
CXT68/80
caaggacGTCTAGACGTCTttgaattcttaGTCTAGACGTCTttgc
CXT185/186
caaggacGTCTAGATGTCTttgaattcttaGTCTAGATGTCTttgc
CXT201/202
caaggacATCATGATATCAttgaattcttaATCATGATATCAttgc
RAD-SMAD
Δpes-10p
GFP
unc-54 3'UTR
BAD-SMAD1
Δpes-10p
GFP
unc-54 3'UTR
BAD-SMAD2
Δpes-10p
GFP
unc-54 3'UTR
B
RAD-SMAD
C
BAD-SMAD1
D
WT
RAD-SMAD
E
lon-2(e678)
RAD-SMAD
F
drag-1(jj4)
RAD-SMAD
G
sma-3(jj3)
RAD-SMAD
H
dbl-1(wk70)
RAD-SMAD
I
WT
BAD-SMAD1
J
relative intensity
0
1
2
3
***
*
***
***
lon-2(e678)
WT
drag-1(jj4)
sma-3(jj3)
dbl-1(wk70)

To test if the expression level or pattern of the RAD-SMAD reporter are regulated by Sma/Mab signaling, we crossed the two integrated RAD-SMAD reporters (*jjIs2433* and *jjIs2436*) into different Sma/Mab pathway mutants and saw similar results with both reporters. As shown in Figure 3.6D-E and 3.6G-H, RAD-SMAD expression in all cell types was down-regulated in *dbl-1* and *sma-3* null mutants, but up-regulated in *lon-2* null mutants. In fact, quantification of the GFP signal intensity in various mutants correlated well with the body sizes of each mutant (Figure 3.6J, compared to the body size of each mutant in Figure 3.2). These results demonstrate that RAD-SMAD positively responds to Sma/Mab signaling.

As shown in Figure 3.6F and 3.6J, RAD-SMAD expression was significantly down-regulated in *drag-1(jj4)* mutants. Furthermore, the degree of reporter down-regulation in *jj4* mutants was smaller than that in *sma-3(jj3)* and *dbl-1(wk70)* mutants, again correlating very well with the body sizes of these mutants. These observations indicate that DRAG-1 is a positive regulator of the Sma/Mab pathway.

3.4 DISSCUSION

3.4.1 DRAG-1 is a cell type-specific modulator of BMP signaling in C. elegans

Previous studies have shown that RGM proteins can enhance BMP, but not TGFβ, signaling in tissue cultures by binding to selected BMP molecules as

well as type I and type II BMP receptors (Babitt et al., 2005, 2006; Samad et al., 2005). However, except for RGMc/Haemojuvelin, there is no evidence showing that RGMa and RGMb/DRAGON are essential for regulating BMP signaling *in vivo* (Babitt et al., 2005, 2006; 2007; Samad et al., 2005; Xia et al., 2008; 2010; Andriopoulos et al., 2009). Our data demonstrate that the sole RGM homolog DRAG-1 is an integral member of the BMP-like Sma/Mab pathway in *C. elegans*: 1) *drag-1* mutants share similar body size and SOSMLP phenotypes as mutants in all the core members of the Sma/Mab pathway, 2) DRAG-1 is a membrane-localized protein that functions at the ligand-receptor level in the Sma/Mab pathway to regulate body size, 3) *drag-1* is expressed and functions in the same cells as the receptors and Smad proteins of the Sma/Mab pathway to regulate body size (in hypodermal cells) and mesoderm patterning (in the M lineage). Both the mutant phenotypes and the reduced level of Sma/Mab pathway reporter (RAD-SMAD) expression in *drag-1* loss-of-function mutants suggest that DRAG-1 positively regulates Sma/Mab signaling in *C. elegans*. DRAG-1 appears to function specifically in the BMP-like Sma/Mab pathway, but not the TGFβ-like dauer pathway, as *drag-1* mutants do not have any defects in dauer formation. The genetic interaction between *drag-1(jj4)* and certain *daf-7* and *daf-1* alleles (Table 3.1) could be due to the fact that both the Sma/Mab pathway and the dauer pathway share the same type II receptor DAF-4 and thus exhibit low levels of crosstalk. Consistent with this notion, previous studies have also shown genetic interactions between *sma-6* and *daf-7* and *daf-1* mutations (Krishna et al., 1999) .

Interestingly, DRAG-1's function in modulating Sma/Mab signaling is cell-type specific. Like other core members of the Sma/Mab pathway, *drag-1* functions in both body size regulation and mesoderm patterning. However, *drag-1* is not expressed in the male tail and *drag-1* mutants do not have any male tail patterning defects. It has been previously suggested that body size and male tail development require different Sma/Mab signaling thresholds, with the male tail requiring a lower level of signaling activity, because a hypomorphic *sma-6(e1482)* allele has small body size but wild-type male tail morphology (Krishna et al., 1999). Our studies on *drag-1* are consistent with this hypothesis, and further demonstrate that *drag-1* functions to augment the level of Sma/Mab signaling activity in the hypodermal cells to ensure proper body size regulation.

3.4.2 RAD-SMAD, a Sma/Mab-responsive reporter in C. elegans

The activity of TGFβ signaling can be monitored by the phosphorylation or nuclear entry of Smads, or the expression of direct TGFβ downstream genes (Schmierer and Hill, 2007; Wrighton et al., 2009). However, there are no existing phospho-Smad antibodies, appropriate Smad::GFP reporters or anti-Smad antibodies, or direct Sma/Mab target genes that allow us to directly monitor the activity of Sma/Mab signaling in *C. elegans*. In this study, we generated a Sma/Mab signaling reporter RAD-SMAD that appears to directly reflect Sma/Mab signaling activity in *C. elegans*: 1) the reporter is active in cells that have been reported to be responsive to Sma/Mab signaling, 2) mutations in the Smad binding sites abolished reporter expression, and 3) the level of RAD-SMAD

reporter expression correlated with the level of Sma/Mab signaling in various Sma/Mab pathway mutants. The RAD-SMAD reporter appears to specifically reflect Sma/Mab signaling, as no reporter expression was observed in dauer animals either in the wild-type background or the *daf-7(e1372)* background. We also noticed that the RAD-SMAD reporter is not expressed in the M lineage, where DRAG-1 and other Sma/Mab pathway components are expressed and function (data not shown). At present, we cannot rule out the possibility that the expression level of RAD-SMAD in the M lineage is too low to detect. An alternative explanation is that expression of Sma/Mab pathway targets in the M lineage requires additional transcriptional input. A likely candidate for the additional transcription factor is the Schnurri (SHN) protein SMA-9 (Liang et al., 2003; Foehr et al., 2006), as previous studies in Drosophila have identified composite Smad-SHN binding sites that contain additional nucleotide sequences in between the two Smad binding sites (Pyrowolakis et al., 2004; Gao et al., 2005). Despite this caveat, the RAD-SMAD reporter will be a useful tool for *C. elegans* researchers wishing to monitor the direct transcriptional output of the Sma/Mab signaling pathway.

In summary, our work demonstrates that DRAG-1 acts in a cell-type specific manner to modulate the BMP-like signaling pathway in *C. elegans*, and establishes a direct link between RGMb/DRAGON proteins and BMP signaling *in vivo*. Because RGM proteins are not present in Drosophila, our work further provides a simple genetic system for mechanistic studies on RGM protein regulation of BMP pathways *in vivo*.

3.5 ACKNOWLEDGEMENTS

We thank the *C. elegans* Genetics Center, Dave Pryune, Shohei Mitani for strains and primers, Aalia AlBarwani, Rachel Fairbank and Amanda Lindy for help with the initial genetic analyses of *drag-1(jj4)*, Ping Wang for help with statistical analysis, and Nirav Amin, Richard Padgett and Mariana Wolfner for helpful discussions and critical comments on the manuscript. This work was supported by NIH R01 GM066953 (to J.L.). Y.P. was a Howard Hughes Undergraduate Research Scholar and a recipient of the CALS Charitable Trust Undergraduate Research Grant at Cornell University.

Chapter 4: *C. elegans* DRAG-1 as a model to study RGM protein function

4.1 Introduction

DRAG-1 belongs to the RGM protein family (Tian et al., 2010). RGMs have an N-terminal signal peptide, a partial von Willebrand type D domain (vWF-type D), which includes a highly conserved autocleavage site, and a C-terminal GPI (Glycosylphosphatidylinositol)-anchor (Monnier et al., 2002;Camus and Lambert, 2007;Corradini et al., 2009;Severyn et al., 2009). Vertebrate RGMa and RGMc proteins have a RGD motif (Figure 4.1). RGM proteins also contain a number of highly conserved cysteine residues (Figure 4.1).

RGM proteins have been shown to function as BMP coreceptors in cell culture and *in vivo*. RGMs are high affinity binding partners of BMP ligands with K_D of several nM (Babitt et al., 2005;Halbrooks et al., 2007;Xia et al., 2007). In cultured cells, RGM proteins selectively bind to a subset of BMP type I and type II receptors (Babitt et al., 2005). Addition of RGMs increased BMP ligand receptor binding (Xia et al., 2007). *In vivo* data suggests that RGMc/HJV functions in modulating BMP6 signaling to regulate the expression of hepcidin (Andriopoulos et al., 2009;Meynard et al., 2009), a key small peptide secreted predominantly from hepatocytes essential for iron homeostasis (Park et al., 2001;Pigeon et al., 2001). RGMb functions in neurite outgrowth and peripheral nerve regeneration by modulating BMP signaling in mice and in cell culture (Ma et al., 2011). The *in vivo* relevance for RGMa functioning in the BMP signaling pathway is not clear.

RGMc/HJV was identified by positional cloning of the locus that is associated with juvenile hemochromatosis (JH) (Papanikolaou et al., 2004). JH is a rare autosomal recessive disease characterized by the early-onset of severe iron overload that affects young patients typically in their first to third decade of life (Camaschella et al., 2002;De Domenico et al., 2008;De Gobbi et al., 2002). JH is also caused by the lack of functional hepcidin (Roetto et al., 2003). Several dozens JH-associated mutations in RGMc/HJV have been identified (Nagayoshi et al., 2008;Robson et al., 2004). An understanding of the mechanism of the disease came from studying some of the missense mutations in cell culture. These studies examined the G99V mutation in that mutated a residue in the RGD motif, the D172E mutation of a key residue for RGM autocleavage, and G320V, which is the predominant disease-causing mutation. G99V mutated RGMc failed to bind to BMP-2, while D172E or G320V mutated RGMc failed to bind to neogenin, an RGM interacting protein that is essential for RGM's function (Kuns-Hashimoto et al., 2008; See Chapter 5). Besides protein-protein interaction, a link has been made between affecting the autocleavage event and protein secretion. Some mutations that lie in or near the cleavage site, D172E, F170C and W191C, as well as G320V abolished the autocleavage event, and these mutated proteins were retained in the endoplasmic reticulum and resulted in low BMP signaling activities (Pagani et al., 2008;Silvestri et al., 2007).

C. elegans has a single RGM protein DRAG-1. DRAG-1 positively modulates BMP-like Sma/Mab signaling in regulating body size and M lineage patterning (Tian et al. 2010). DRAG-1 is a membrane-associated protein that

functions at the ligand-receptor level to modulate the Sma/Mab pathway in a cell-type-specific manner. Study of DRAG-1 provides a direct link between RGM proteins and BMP signaling *in vivo* and a simple and genetically tractable system for mechanistic studies of RGM protein regulation of BMP pathways. In this Chapter, I first show that vertebrate RGMs can replace DRAG-1 in *C. elegans* Sma/Mab signaling, suggesting a high degree of functional conservation in RGM and BMP signaling in *C. elegans*. I then used DRAG-1 as a qualitative and quantitative *in vivo* model to test three JH mutations that were studied previously in cell cultures (Kuns-Hashimoto et al., 2008). Consistent with previous results, all three mutations affect DRAG-1 function in Sma/Mab signaling, but they do so to different degrees. Lastly, I demonstrated DRAG-1 physical interaction with Sma/Mab pathway ligand and receptors. This suggests that DRAG-1 may function by simultaneously interacting with both the ligand and receptors.

4.2 Material and Methods

4.2.1 C. elegans strains

Strains were kept under normal conditions, described by Brenner (1974). Analyses were performed at 20°C, unless otherwise noted. The following mutations and integrated transgenes were used: Linkage group I (LGI): *drag-1 (jj4)*; LGX: *sma-9(cc604)*.

4.2.2 Plasmid constructs and transgenic lines

Constructs for testing vertebrate RGM function in C. elegans

pCXT14: *drag-1p* (4kb of *drag-1* upstream sequence (-3977 to -1), sequences from immediately downstream of the ATG till the end of the first intron of *drag-1* (4 to 1123))::full length mouse *Rgmb* coding sequence::*unc-54* 3'UTR.

pCXT161: *drag-1p* (same as above)::*drag-1* signal sequence:: *mRgmb* mature region::*drag-1* C- signal sequence::*unc-54* 3'UTR, has aa23–aa375 of DRAG-1 replaced by aa59–aa414 of mouse Rgmb in pCXT15 (Tian et al., 2010).

pCXT185: *drag-1p* (same as above)::*drag-1* signal sequence:: *hHJV* mature region::*drag-1* C- signal sequence::*unc-54* 3'UTR, has aa23–aa375 of DRAG-1 replaced by aa38–aa398 of human RGMc/HJV in pCXT15 (Tian et al., 2010).

Constructs for interaction analysis of DRAG-1

pCXT239: pCMV::signal peptide::FLAG-*drag-1* mature region, with a FLAG tag in the N- and a stop codon C-terminal to aa22-aa360 of DRAG-1 in pSecTag.

pCXT241: pCMV::signal sequence::*dbl-1* prodomain::FLAG-*dbl-1* mature region, has aa32–stop codon of DBL-1 in pSecTag construct, FLAG tag is inserted between aa238 and aa239.

pJKL964: pCMV::signal sequence::c-Myc-His-*sma-6* full length, has aa27–stop of SMA-6 in pSecTag construct with c-Myc-His tagged on the N terminus after the signal sequence.

pCXT230: pCMV::signal sequence::Myc-His-*daf-4* EXD, TM and partial ICD, has aa33–aa468 of DAF-4 in pSecTag construct with c-Myc-His tagged on the N terminus after the signal sequence.

pJKL962: pCMV::signal sequence::Myc-His-*drag-1* mature domain, with a c-Myc-His tag in the N- and a stop codon C-terminal to aa22-aa360 of DRAG-1 in pSecTag construct.

Generating transgenic animals

Transgenic animals (except for the ones generated by MosSCI) were generated using the plasmids pJKL449 (*myo-2p::gfp::unc-54* 3' UTR, Jiang et al., 2009) as markers.

4.2.3 Body size measurement

Hermaphrodite animals at the gravid adult stage were collected and treated with hypochlorite. The resulting embryos were allowed to hatch in M9 buffer at 16°C. Synchronized L1s were plated onto NGM plates and allowed to grow for 72h before they were washed off the plates, treated with 0.3% sodium azide, and mounted onto 2% agarose pads. Images of the worms were taken on a Leica DMRA2 compound microscope with a Hamamatsu Orca-ER Camera using the Openlab software (Improvision). Body length of animals whose vulva development is at early Charismas tree stage was measured using Openlab software. Subsequent statistics analyses were performed using Microsoft Excel.

4.2.4 MosSCI technique

Insertion lines were generated using direct gonadal injection method using *ttTi5605* Mos site and *pglh-2* driving Mos transposase following protocols in Frokjaer-Jensen et al. (2008).

Constructs used in this study:

pCXT208: pJKL849 (Tian et al., 2010) was modified to delete sequences encoding intron 3, because the repetitive sequences make it extremely hard for cloning. The new construct is pCXT183. 3.9kb *drag-1p*::*drag-1 genomic sequence without 3*rd *intron*::*gfp*::1.7*kb drag-1* 3' UTR was moved from pCXT183 with restriction enzymes HpaI and StuI to pCFJ151 digested by XhoI and blunted (Frokjaer-Jensen et al., 2008).

pCXT221: G68V (ggt to gtt) mutation was made in pCXT183. The genomic sequence was then moved in the same way as shown above to pCFJ151.

pCXT222: D117E (gat to gaa) mutation was made in pCXT183. The genomic sequence was then moved in the same way as shown above to pCFJ151.

pCXT223: G272V (gga to gta) mutation was made in pCXT183. The genomic sequence was then moved in the same way as shown above to pCFJ151.

PCR verifications for insertions:

PCR pair to verify the N terminus of the insertion:

CXT310: gcgggatcatttcttactag. Sense, anneals to genomic sequence till 0bp upstream of Mos insertion site.

CXT226: cctgaatttgtaaatactcttc. Reverse, anneals to *drag-1* promoter sequence in pCXT208.

PCR pair to verify the C terminus of the insertion.

CXT312: caaggacttggataaattggc. Sense, anneals to C-terminal homology arm in pCFJ151.

CXT311: gtgtatctgcattaaccaata. Reverse, anneals to genomic sequence till 0bp downstream of Mos insertion site.

4.2.5 Co-immunoprecipitation

Medium was collected from HEK293T cells that were transiently transfected by constructs carrying FLAG-tagged or Myc-tagged DRAG-1 mature region (pCXT239 or pJKL962 respectively) and FLAG-tagged DBL-1 mature domain (pCXT241) for five days. Cell lysates were harvested 48h post transfection from HEK293T cells that are transiently transfected by constructs carrying Myc-tagged SMA-6 (pJKL964) and Myc-tagged DAF-4 (pCXT230). Anti-c-Myc conjugated sepharose or anti-FLAG M2 conjugated sepharose was used to incubate with media or cell lysates expressing one protein for 4 hour at 4°C, and then with media or cell lysates expressing a second protein for 4 hours. The beads were then subjected to 5 time wash with 50 mM Tris, pH 8.0, 150 mM NaCl, and 1% Triton X-100. The presence of DRAG-1 in the immunoprecipitates

was detected by western blotting using mouse anti-FLAG (M2) antibodies (4mg/ml, 1:10000) or anti-c-Myc (9E10) antibodies (2mg/ml, 1:5000).

4.3 Results

4.3.1 Vertebrate RGMs are functional in C. elegans

DRAG-1 shares high degree of homology to the vertebrate RGM proteins in its mature protein domain but not the N- and C- terminal signal sequences (Figure 4.1). DRAG-1 functions in a similar fashion to its vertebrate homologs in modulating the BMP-like Sma/Mab signaling (Tian et al., 2010). I therefore asked if *C. elegans* DRAG-1 provides a good system for studying RGM protein function. I started by testing if vertebrate RGMs are functional in *C. elegans*. As in Chapter 3, I will refer to the suppression of the *sma-9(cc604)* M lineage phenotype by *jj4* as the SOSMLP phenotype of *jj4* in the remainder of Chapter 4 and 5. As a control, a construct carrying *drag-1* promoter driving *drag-1* cDNA is functional to rescue both the small body size and the SOSMLP phenotypes (Table 4.1). However, the rescue of SOSMLP by this cDNA construct is only 17.9%, which is incomplete rescue because the rescue efficiency of the *drag-1* genomic construct is 100% (Table 4.1). The rescue of small body size is much better by this cDNA construct (Figure 4.2). This could be due to the lack of an M lineage enhancer in the promoter; this may also be due to the general lower expression level of *drag-1* in the cDNA construct compared to the genomic construct and to the possibility that different processes have different sensitivity to Sma/Mab signaling. To

examine whether vertebrate RGM could function similarly to *drag-1* in *C. elegans*, I replaced the drag-1 cDNA with the mouse Rgmb cDNA. However, this construct failed to rescue in both the body size rescue and the SOSMLP rescue assays (Table 4.1 and Figure 4.2). Because of the lack of homology in the N- and C-terminal signal sequences, I reasoned that *Rgmb* signal sequences were not recognized by *C. elegans* system and that the Rgmb protein might not be properly localized in *C. elegans*. Therefore, I generated a hybrid *drag-1-Rgmb* gene with *drag-1* N- and C- terminal signal sequences and mouse Rgmb mature protein sequences under the control of *drag-1* promoter. The hybrid *drag-1-Rgmb* transgene was able to rescue both the small body size and the SOSMLP of *sma-9* in a comparable degree as *drag-1* cDNA itself (Table 4.1 and Figure 4.2). I also tested a hybrid *drag-1*-human *RGMc/HJV* cDNA construct and showed that it functions as well as *drag-1* cDNA control in rescuing the SOSMLP (Table 4.1). I am currently testing RGMc/HJV construct using the body size assay. In summary, these results indicate that after proper processing, vertebrate RGMs are functional in *C. elegans* to a similar degree as *drag-1* cDNA construct. Therefore, *drag-1* could potentially be a good model to study human RGM function *in vivo*.

Table 4.1: Rescue of the SOSMLP by *drag-1* with various constructs.

Transgenes in *sma-9(cc604); drag-1(jj4)*	The extent of suppression
drag-1::gfp genomic	100% (N=43)
drag-1 cDNA	17.9% (N=106)
Rgmb cDNA	0.9% (N=107)
drag-1-Rgmb hybrid cDNA	14.6% (N=82)
drag-1-HJV hybrid cDNA	23.4% (N=143)
Mos-SCI *drag-1:gfp*	98.8% (N=498)
Mos-SCI G68V in *drag-1::gfp*	20.7% (N=554)
Mos-SCI D117E in *drag-1::gfp*	99.2% (N=515)
Mos-SCI G272V in *drag-1::gfp*	5.1% (N=414)

Figure 4.1: **Comparison between DRAG-1 and vertebrate RGM proteins.**

Alignment between DRAG-1 (the bottom line) and its human RGM homologs RGMa (GenBank AAI51133.1), RGMb (GenBank NP_001012779.2) and RGMc (GenBank NP_998818.1). The purple shaded area is the N- and C- terminal signal sequences predicted for *drag-1* sequence. The green shaded area represents the highly conserved vWF type D domain. The RGD motif (boxed in red) is present in RGMa and RGMc, but not in RGMb and DRAG-1. The black vertical dotted line represents the autocleavage site that is conserved in all four RGM proteins. Red underlines represent highly conserved Cysteine residues. Blue underline represents the motif that is sufficient for DRAG-1-UNC-40 interaction (See Chapter 6). The arrows point to the mutated residues in human JH disease with molecular lesion shown on top of the arrows.

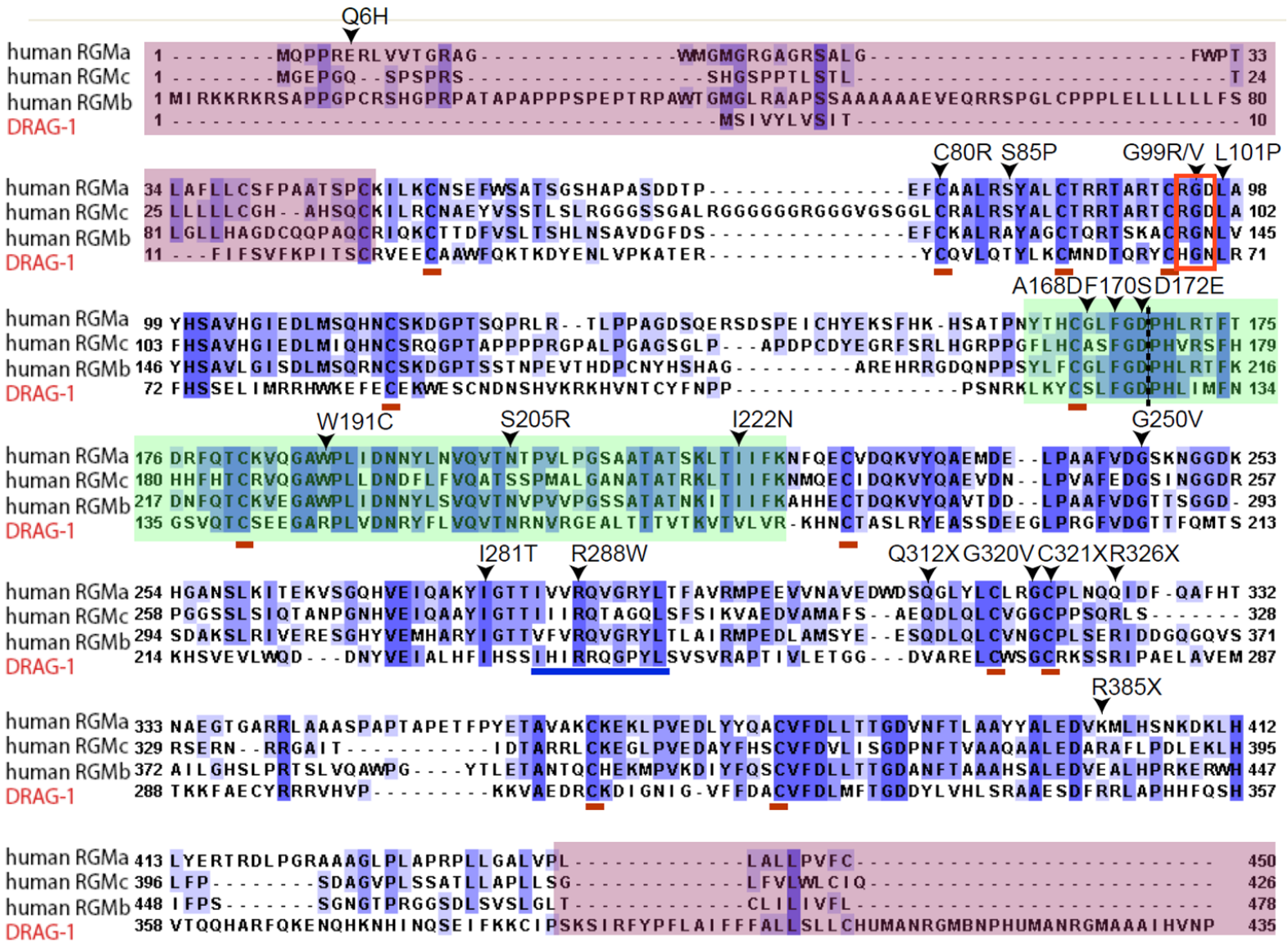
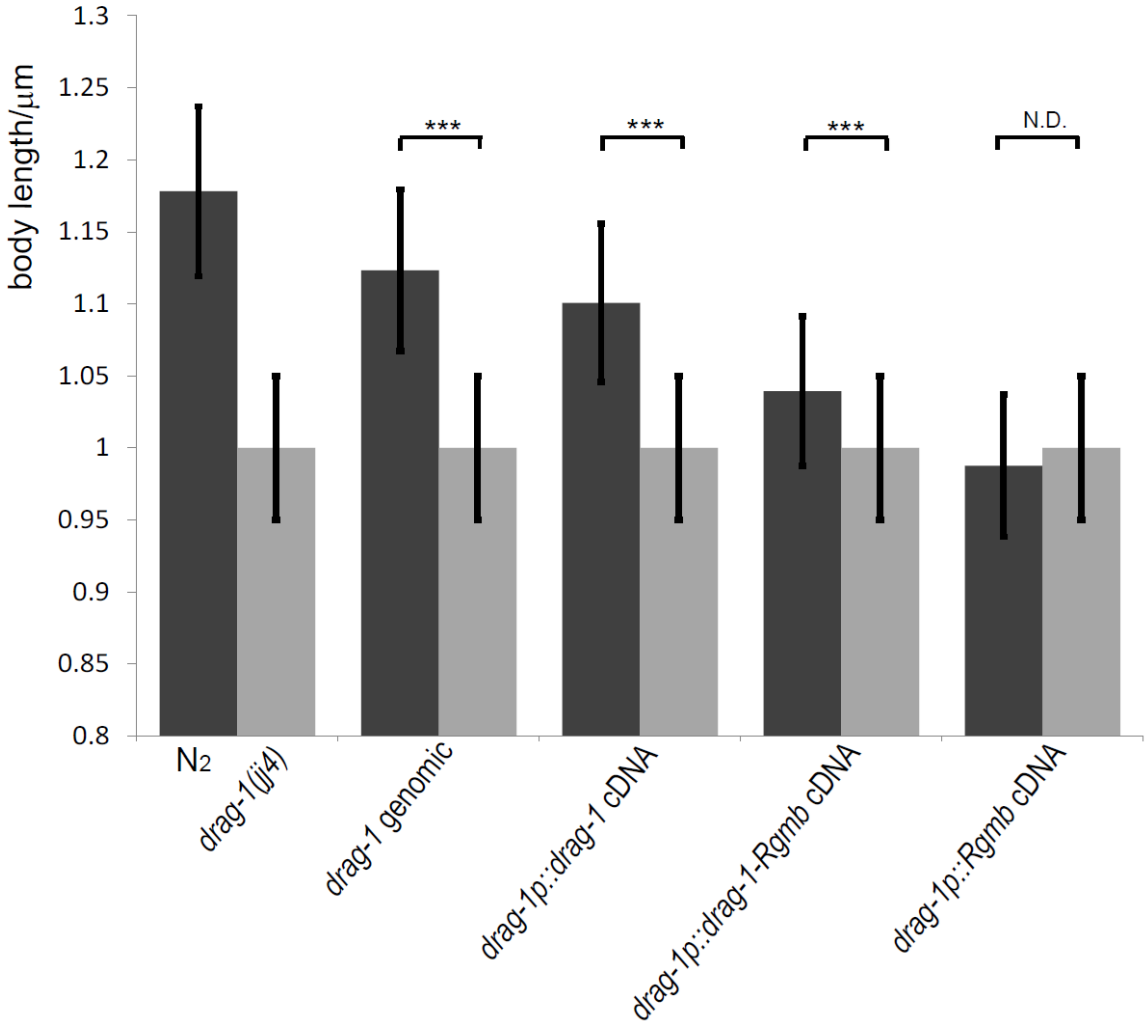

Q6H
C80R
S85P
G99R/V
L101P
A168D
F170S
D172E
W191C
S205R
I222N
G250V
I281T
R288W
Q312X
G320V
C321X
R326X
R385X
human RGMa
human RGMc
human RGMb
DRAG-1

Figure 4.2: Body size rescue of *drag-1(jj4).*

Body size was measured in adult worms at 72 hours post-plating. Transgenic worms were distinguished from non-transgenic ones by the presence of *myo-2::gfp* in the former. The lengths of *drag-1(jj4)* and non-transgenic worms of each group were normalized to 1. The relative lengths of wild-type and various transgenic animals compared with their corresponding non-transgenic controls are presented. Between 12 and 44 animals were measured for each genotype. Error bars represent 95% confidence intervals for the mean of relative body length. The significance of difference between transgenic and the corresponding non-transgenic animals was statistically analyzed via Student's t-test. ***, $P<0.0001$; N.D., not significantly different.


body length/µm
1.3
1.25
1.2
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
***
***
***
N.D.
N2
drag-1(jj4)
drag-1 genomic
drag-1p::drag-1 cDNA
drag-1p::drag-1-Rgmb cDNA
drag-1p::Rgmb cDNA

4.3.2 Assessment of human HJV mutations using C. elegans DRAG-1 as a model

Several dozen human JH disease mutations are identified in *RGMc/HJV* (Nagayoshi et al., 2008;Robson et al., 2004). Missense mutations associated with JH disease are shown in Figure 4.1. Most of the residues carrying whose mutation in human causes JH disease are conserved in DRAG-1 (Figure 4.1). Several of the missense mutations have been characterized *in vitro* for their function (Kuns-Hashimoto et al., 2008). However, there haven't been *in vivo* studies to assess their consequences. To provide a qualitative and quantitative assessment of these human mutations, I used DRAG-1 as an *in vivo* model to test human mutations in three conserved DRAG-1 residues: G68 (G99 in human HJV) that is located in an RGD motif that is present in RGMa, HJV, but not in RGMb or DRAG-1; D117 (D172 in human HJV) that is located in the critical residue for the conserved "GDPH" motif for RGM autocleavage; G272 (G320 in human HJV) that is the most frequently mutated residue in JH patients (Figure 4.1). I generated the G68V, D117E, and G272V mutations in functional *drag-1::gfp* genomic constructs (Tian et al., 2011). I used the newly developed Mos-mediated single copy insertion (Mos-SCI, (Frokjaer-Jensen et al., 2008)) technique to insert the mutation-carrying constructs as well as a wild type *drag-1::gfp* construct into the same genomic locus as single copy. Two insertion lines were generated for each of the wild type, D117E, and G272V constructs and one insertion line was generated for G68V construct. The insertions were PCR verified (See Material and Methods). Western blotting showed that proteins from

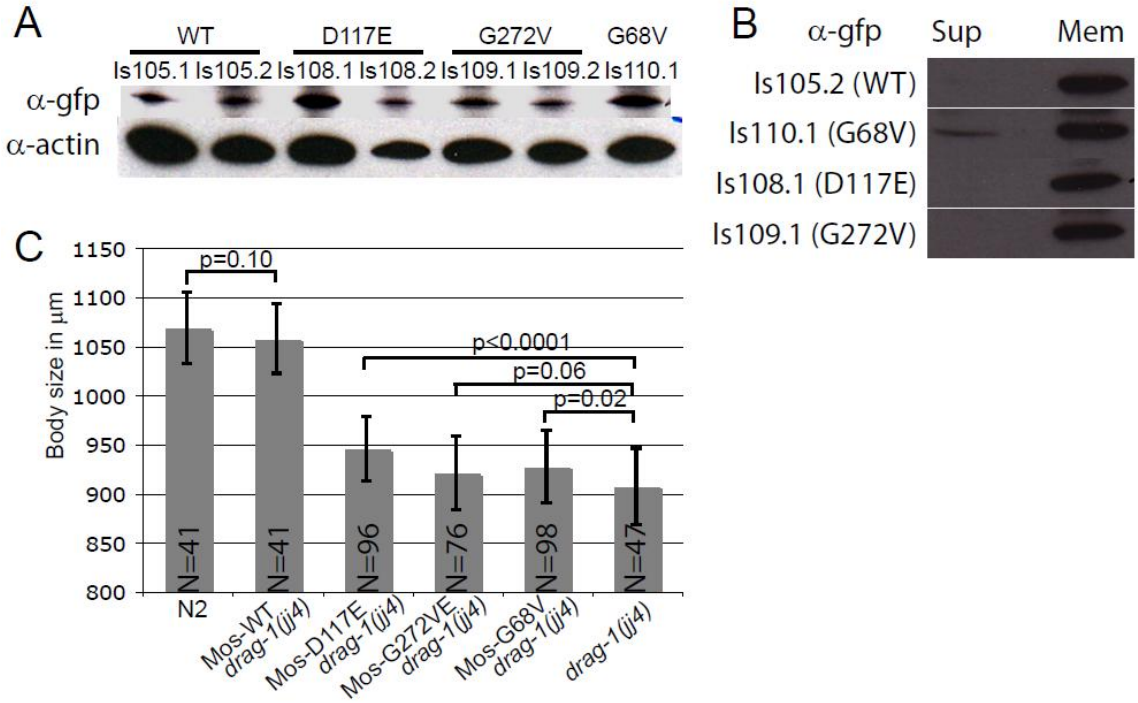
each insertion line are expressed at a similar level and the same size (Figure 4.3A). Assuming that each line has one copy of transgene insertion per chromosome, the point mutations do not affect the protein stability. Fractionation experiments showed that neither wild type DRAG-1::GFP nor the three mutated forms of DRAG-1::GFP are mainly or totally localized to the membrane fraction, with very little of G68V mutated DRAG-1::GFP became soluble (Tian et al., 2010, Figure 4.3B). Thus, the mutations do not affect DRAG-1 protein membrane association.

The insertions were then crossed into the *drag-1(jj4)* background to test for rescue of small body size phenotype and into the *drag-1(jj4); sma-9(cc604)* background to test for the rescue of the SOSMLP. The results were similar in both assays between the two independent inserted lines for each of wild type, D117E, and G272V (data not shown). Results were thus averaged; they are summarized in Table 4.1 and Figure 4.3C. In both assays, wild type *drag-1::gfp* is fully functional: it rescues the small body size to wild type levels and rescues the SOSMLP close to 100%. Among G68V, D117E, and G272V mutated *drag-1::gfp* transgenes, G272V is the least functional: it failed to rescue the small body size of *drag-1(jj4)* mutant, and it only rescued 5.1% of the SOSMLP. G68V is moderately functional: it significantly rescued the small body size phenotype and rescued the SOSMLP in 20.7% of the worms. Interestingly, D117E significantly rescued the small body size but not nearly to the wild type level and fully rescued the SOSMLP. One explanation for this result is that body size is likely more sensitive to DRAG-1 protein function and Sma/Mab signaling activity than M

lineage patterning. Together, these results suggest that the three mutations affect DRAG-1 protein function to different levels, with G272V being the most impaired and with D117E being the least impaired.

Figure 4.3: JH disease mutations in DRAG-1 affect DRAG-1 function to different degrees.

(A) Western blotting showing similar expression level of different DRAG-1::GFP forms in Mos-SCI lines with actin as a loading control (see Materials and methods). (B) Wild type and mutant forms of DRAG-1::GFP are associated with the membrane fraction (see Materials and methods). (C) Body length of wild-type (N2) and various mutant worms at 72 hours post-plating (see Materials and methods). Error bars represent standard deviation. The significance of difference between double mutants and the corresponding single mutants was statistically analyzed via Student's t-test. P values are shown.


A
WT
D117E
G272V
G68V
Is105.1 Is105.2 Is108.1 Is108.2 Is109.1 Is109.2 Is110.1
α-gfp
α-actin
B
α-gfp
Sup
Mem
Is105.2 (WT)
Is110.1 (G68V)
Is108.1 (D117E)
Is109.1 (G272V)
C
Body size in µm
1150
1100
1050
1000
950
900
850
800
p=0.10
p<0.0001
p=0.06
p=0.02
N=41
N=41
N=96
N=76
N=98
N=47
N2
Mos-WT drag-1(jj4)
Mos-D117E drag-1(jj4)
Mos-G272VE drag-1(jj4)
Mos-G68V drag-1(jj4)
drag-1(jj4)

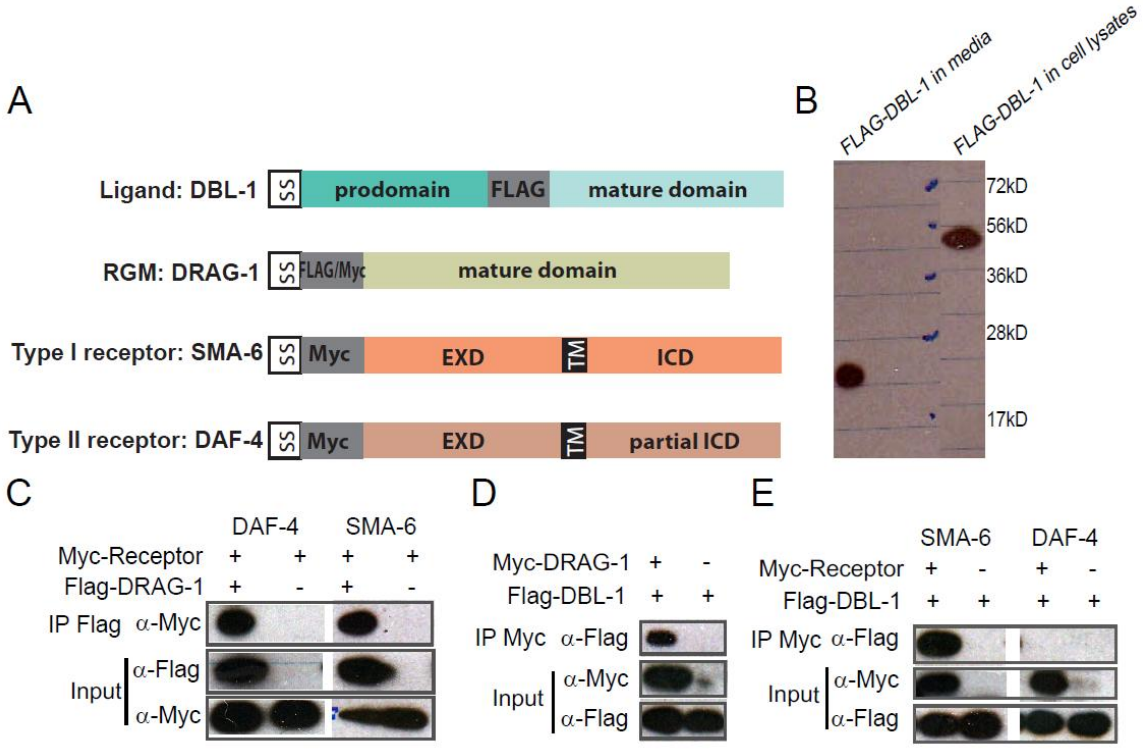
4.3.3 DRAG-1 physically interacts with the Sma/Mab ligand DBL-1, type I receptor SMA-6, and type II receptor DAF-4.

Vertebrate RGM proteins physically associate with the BMP receptors and directly bind to the BMP ligand with high affinity in mammalian cell culture. I therefore tested whether DRAG-1 binds to Sma/Mab ligand and receptors. Flag-tagged DRAG-1 mature domain and Myc- tagged SMA-6 mature region or Myc-tagged DAF-4 mature region that lacks aa 439-end were transfected into mammalian HEK293T cells (Figure 4.4A). Immunopreciptation (IP) using anti-FLAG successfully pulled down both Myc-SMA-6 and Myc-DAF-4 proteins, suggesting that DRAG-1 physically interacts with both SMA-6 and DAF-4 (Figure 4.4C). DBL-1, like the other TGF-β family ligands, goes through proprotein convertase cleavage between its prodomain and mature domain to release the active mature domain. Therefore, I generated a construct expressing FLAG-mature DBL-1 by inserting the FLAG tag before the DBL-1 mature domain and after the proprotein cleavage site in DBL-1 prodomain (Figure 4.4A). The FLAG-mature DBL-1 transfected HEK293T media contained the mature FLAG-DBL-1 of the expected size, while the cell lysate contained the full length unprocessed form of DBL-1 prodomain-FLAG-mature DBL-1 (Figure 4.4B). Anti-c-Myc conjugated sepharose was first incubated with DRAG-1 that was expressed as a Myc-DRAG-1 form in HEK293T cells, and then with HEK293T media containing DBL-1. IP pulled down FLAG-mature DBL-1 (Figure 4.4D). Thus DRAG-1 also physically interacts with the ligand DBL-1. As a positive control, type I receptor SMA-6 physically interacted with the ligand DBL-1 while type II receptor DAF-4

failed to bind to DBL-1 (Figure 4.4E). This is consistent with previous finding that BMP ligands have higher affinity to the type I receptors than type II receptors (Attisano et al., 1994;Massague, 1992;Massague, 1998). These results and previous studies (Tian et al., 2010; Chapter 3) support the hypothesis that DRAG-1 is likely a Sma/Mab pathway co-receptor. For confirmation, however, I am currently examining all of these pairs of interactions with reciprocal IPs.

Figure 4.4: DRAG-1 physically interacts to the Sma/Mab receptors SMA-6 and DAF-1 and the ligand DBL-1.

(A) Schematics of all the constructs used in this study, with ss representing the N-terminal signal sequence, TM representing the transmembrane domain, EXD representing the extracellular domain, and ICD representing the intracellular domain. (B) Western blotting of FLAG-DBL-1 in media and cell lysates showed different molecular sizes indicative of protein processing. Co-IP experiment using the above constructs showed that SMA-6 and DAF-4 (B) and mature domain of DBL-1 (C) all bind to DRAG-1 mature domain. As controls, SMA-6, but not DAF-4, binds to DBL-1 (D).


A
Ligand: DBL-1
SS
prodomain
FLAG
mature domain
RGM: DRAG-1
SS
FLAG/Myc
mature domain
Type I receptor: SMA-6
SS
Myc
EXD
TM
ICD
Type II receptor: DAF-4
SS
Myc
EXD
TM
partial ICD
B
FLAG-DBL-1 in media
FLAG-DBL-1 in cell lysates
72kD
56kD
36kD
28kD
17kD
C
DAF-4
SMA-6
Myc-Receptor + + + +
Flag-DRAG-1 + - + -
IP Flag α-Myc
Input α-Flag
α-Myc
D
Myc-DRAG-1 + -
Flag-DBL-1 + +
IP Myc α-Flag
Input α-Myc
α-Flag
E
SMA-6
DAF-4
Myc-Receptor + - + -
Flag-DBL-1 + + + +
IP Myc α-Flag
Input α-Myc
α-Flag

4.4 Discussion

In this study, I established that DRAG-1 and vertebrate RGMs share functional similarities. DRAG-1 has cysteine residues at positions identical to vertebrate RGMs, suggesting that similar pairs of disulfide bonds might form between comparable cysteine residues, leading to a highly similar protein structure. Human JH disease missense mutations are located in the residues that are conserved in DRAG-1. Furthermore, examination of three of the human mutations introduced into DRAG-1 in *C. elegans* showed that the mutations have quantitatively and qualitatively different impacts on DRAG-1 protein function. The most prevalent human mutation, corresponding to G272V in DRAG-1, has the most abolishment on DRAG-1 function. This mutation (G320V in human RGMc/HJV) abolished RGMc interaction with neogenin protein in cell culture (See Chapter 5, Zhang et al., 2009). This is consistent with the notion that neogenin is involved in the BMP signaling (Zhang et al., 2009;Zhou et al., 2010; Lee et al., 2010). As will be discussed in Chapter 5 of this thesis, I showed that *C. elegans* neogenin UNC-40 functions in modulating Sma/Mab signaling. Therefore, I will test if *C. elegans* G272V, the corresponding RGMc G320V mutation in DRAG-1, disrupts DRAG-1 interaction to UNC-40 and if the disruption of this interaction affects Sma/Mab signaling.

I also established that *C. elegans* DRAG-1 physically interacts with the Sma/Mab ligand DBL-1 and both type I receptor SMA-6 and type II receptor DAF-4 in HEK293T cells. Similar results were published in mammalian cultured cells showing that all three RGMs were able to directly bind to BMPs and

physically associate with both type I and type II receptors (Babitt et al., 2005;Babitt et al., 2006). Because of the profound similarity in protein sequences between DRAG-1 and its vertebrate homologs, DRAG-1 may interact with the Sma/Mab ligand and receptors through some conserved motifs, rendering DRAG-1 suitable for mapping out critical motifs for the interactions.

Therefore, *C. elegans* has a highly conserved BMP-like Sma/Mab signaling system that includes a highly conserved role for the RGM protein DRAG-1, and together with the worm's easily distinguishable and nonessential Sma/Mab pathway phenotypes, which makes *C. elegans* a great tool to study BMP signaling and RGM protein function. The recent availability of the single copy insertion technique, Mos-SCI, makes *C. elegans* even more powerful to quantitatively assay for protein function (Frokjaer-Jensen et al., 2008). In this study, I used transgenes introduced by the Mos-SCI method to compare the function of DRAG-1 proteins carrying different point mutations using two assays, the rescue of small body size of *drag-1(jj4)*, and the rescue of SOSMLP. Results from both assays suggest that G272V mutation is the least functional, G68V mutation is intermediate functional, and D117E mutation is the most functional. An unexpected and yet interesting finding come out from these comparisons is that I found that different tissues respond to differently to the mutations, which may be due to different levels of Sma/Mab signaling. This wouldn't have been identifiable by conventional transgenic techniques, which cannot be used for quantitative comparison of transgenes because of high copy numbers of transgenes that lead to severe overexpression. Specifically, D117E mutation

couldn’t rescue the small body size phenotype based on a moderate cutoff ($p<0.1\%$) but fully rescued the *sma-9* M lineage suppression. One reasonable explanation from this study is that M lineage patterning requires lower level of Sma/Mab signaling than body size regulation.

The next step of the project is to determine why the mutations, especially the G272V mutation, affect DRAG-1 protein function. Two directions will be the main focus: to determine if the mutated DRAG-1 proteins fail to bind to its crucial partners, and to investigate if the mutations affect DRAG-1 subcellular localization (See Chapter 6).

4.5 Acknowledgement

I want to thank Fenghua Hu for help with the co-IP assay, Herong Shi for preparing constructs and for PCR genotyping.

Chapter 5: UNC-40/neogenin functions independently of netrin signaling in modulating the BMP-like Sma/Mab signaling in *C. elegans*

5.1 INTRODUCTION

Bone morphogenetic proteins (BMPs) belong to the transforming growth factor-β (TGFβ) superfamily of ligands. BMP signaling pathway plays important roles in development and disease (Wu and Hill, 2009). The BMP signal transduction mode is relatively simple (Shi and Massague, 2003). The signal activation starts with the formation of the BMP ligand-type I,II receptor complex. Once in close proximity, the constitutively active type II receptor phosphorylates the type I receptor, which then phosphorylates the receptor-regulated Smads (R-Smads). These activated R-Smads then complex with common-mediator Smads (co-Smads) and enter the nucleus. The R-Smad-co-Smad complex is then joined by transcription cofactors in regulation of downstream gene expression. Multiple levels of regulation, including extracellular regulation, are critical to ensure proper spatiotemporal control of BMP signaling (Balemans and Van Hul, 2002;Massague and Chen, 2000;Moustakas and Heldin, 2009;Umulis et al., 2009).

The repulsive guidance molecule (RGM) protein family, including RGMa, RGMb, and RGMc (hemojuvelin or HJV), as the name suggests, was initially originally identified because of a role in axon guidance (Monnier et al., 2002). Since then, members of the family have also been shown to function as BMP co-recetors by *in vitro* and *in vivo* studies in vertebrates (Babitt et al., 2005;Babitt et

al., 2006;Samad et al., 2005). All three RGM proteins physically interact with selective BMP ligands and receptors to enhance BMP signaling in cell cultures (Babitt et al., 2005;Babitt et al., 2006;Samad et al., 2005;Xia et al., 2007;Xia et al., 2008). In mice, RGMc acts as the BMP6 co-receptor in regulating expression of hepcidin, a peptide that regulates iron metabolism (Park et al., 2001;Pigeon et al., 2001;Andriopoulos et al., 2009;Babitt et al., 2007;Meynard et al., 2009). Mutations in RGMc/HJV in humans cause juvenile hemochromatosis, an iron overload disease with a reduction in hepatic hepcidin expression (Papanikolaou et al., 2004a;Papanikolaou et al., 2004b). The single RGM family member in *C. elegans*, DRAG-1, functions similarly as a co-receptor of the BMP-like Sma/Mab pathway (Tian et al., 2010).

Neogenin, a member of the DCC (deleted in colorectal cancer) family, regulates neuronal axon guidance by serving as a receptor for the guidance cue netrin (Keino-Masu et al., 1996) as being a receptor for the repulsive cue RGM (Rajagopalan et al., 2004;Cole et al., 2007). In addition to those roles, recent *in vitro* and *in vivo* evidence has shown that neogenin also modulates BMP signaling. Neogenin-RGM interaction is required for the activation of hepcidin expression downstream of BMP signaling in mice and cultured cells (Zhang et al., 2009;Lee et al., 2010). Neogenin also positively regulates BMP-induced endochondral bone formation in mice (Zhou et al., 2010). In cell cultures, neogenin has been shown to recruit BMP receptors to lipid rafts where the signaling occurs (Zhou et al., 2010). Neogenin has also been shown play negative roles in modulating BMP signaling in cell cultures: it facilitates the

secretion of soluble RGMc, a negative regulator of BMP signaling (Zhang et al., 2005;Zhang et al., 2007); neogenin-BMP ligand interaction also activates RhoA signaling events that antagonize Smad activation (Hagihara et al., 2011). However, none of the negative roles of neogenin in modulating BMP signaling were identified *in vivo*. Therefore, an *in vivo* genetic model is valuable for understanding how neogenin modulates BMP signaling.

C .elegans contains a single protein, UNC-40, that is homologous to both DCC and neogenin (Hong et al,. 1996). In this study, I examined the role of UNC-40 in modulating the BMP-like Sma/Mab pathway. There are two BMP-related pathways in *C. elegans*, a dauer pathway that controls dauer formation and a Sma/Mab pathway that regulates body size, male tail formation, mesoderm development and innate immunity (Foehr et al., 2006;Partridge et al., 2010;Savage-Dunn, 2005). The ligand of the Sma/Mab pathway is a BMP-like molecule DBL-1 (Morita et al., 1999;Suzuki et al., 1999). Additional members of this pathway include the type I receptor SMA-6 (Krishna et al., 1999), type II receptor DAF-4 (Estevez et al., 1993), and the Smad proteins, SMA-2, SMA-3 and SMA-4 (Savage et al., 1996). Loss-of-function mutations in any of these pathway members cause small body size and male tail sensory ray formation defects.

The Sma/Mab pathway also plays a role in the *C. elegans* postembryonic mesoderm development. The postembryonic mesodermal M lineage arises from a single precursor cell, the M mesoblast. During hermaphrodite postembryonic

development, the M cell divides to produce a group of dorsal cells that include the macrophage-like coelomocytes (CCs) and striated bodywall muscles (BWMs), and a group of ventral cells that includes BWMs and the sex muscle precursor cells, the sex myoblasts (SMs; Fig. 1C; Sulston and Horvitz, 1977). We have previously shown that mutations in the schnurri homolog *sma-9* lead to a dorsal-to-ventral fate transformation in the M lineage (Foehr et al., 2006;Foehr and Liu, 2008). Furthermore, mutations in the core components of the Sma/Mab pathway suppress the dorsoventral patterning defects of *sma-9* mutants, suggesting that the Sma/Mab signaling is antagonized by SMA-9 to pattern the postembryonic mesoderm along the dorsoventral axis.

In this study, I identified UNC-40 as a positive modulator of the Sma/Mab pathway. I showed that the role of UNC-40 in the Sma/Mab pathway is independent of the netrin cue UNC-6 and the netrin receptor UNC-5. I demonstrated that UNC-40 binds to and synergizes with the RGM protein DRAG-1 via its extracellular domain to mediate Sma/Mab signaling. Surprisingly, the intracellular domain of UNC-40 is not required in this process.

5.2 MATERIALS AND METHODS

5.2.1 C. elegans strains

Strain handling is the same as in Chapter 2, 3, and 4. The following mutations and integrated transgenes were used: Linkage group I (LGI): *drag-1(jj4)*; *unc-40(e1430)*; *unc-40(tr115)*; *unc-40(ev457)*; *unc-40(ev546)*; *unc-*

40(e271); *unc-40(ur304)*; LGII: *sma-6(jj6)*; LGIII: *lon-1(e185)*, *ccIs4438(intrinsic CC::gfp);* LGIV: *unc-5(e53)*; LGV: *dbl-1(wk70)*; LGX: *lon-2(e678); sma-9(cc604)*; *unc-6(ev400)*.

5.2.2 Constructs for interaction analysis of UNC-40

pCXT239: pCMV::signal peptide::FLAG-*drag-1* mature region, with a FLAG tag in the N- and a stop codon C-terminal to aa22-aa360 of DRAG-1 in pSecTag.

pCXT245: pCMV::signal peptide::c-MycHis-*unc-40* FNIII 5-6, with a cMycHis tags in the N- and a stop codon C-terminal to aa852-aa1081 of UNC-40 in pSecTag.

5.2.3 Body size measurement

Hermaphrodite animals at the gravid adult stage were collected and treated with hypochlorite. The resulting embryos were allowed to hatch in M9 buffer at 16°C. Synchronized L1s were plated onto NGM plates and allowed to grow for 48h before they were washed off the plates, treated with 0.3% sodium azide, and mounted onto 2% agarose pads. Image taking and data analysis are the same as in Chapter 4.

5.2.4 Immunofluorescence staining

Animal fixation, immunostaining, microscopy and image analysis were performed as described previously (Tian et al., 2011). Guinea pig anti-FOZI-1 (1:200; Amin et al., 2007), mouse monoclonal anti-AJM-1 (clone MH27, 1:100), rat anti-MLS-2 (1:200; Jiang et al., 2005) and goat anti-GFP (Rockland

Immunochemicals; 1:1000) were used. All secondary antibodies were from Jackson ImmunoResearch Laboratories and used at a dilution of 1:50 to 1:200.

5.2.5 RAD-SMAD reporter assay

Hermaphrodites carrying the RAD-SMAD reporters (jjIs2437) were treated with hypochlorite. The resulting embryos were allowed to hatch in M9 buffer at 16°C. Synchronized L1s were plated onto NGM plates and allowed to grow for 48 hours. The resulting L3 animals were anesthetized. Images of the hypodermal cells were taken at 40x magnification on a Leica DMRA2 compound microscope. Fluoresence intensity of the nuclei and background was measured in pairs using the Openlab software. Five nuclei/background pairs were measured for each worm. The value of nuclei signal minus background signal was averaged for each worm. The middle 60% of the means (discarded 20% of the highest and 20% of the lowest means) were averaged to generate the value of signal intensity for a given genotype. Standard deviation was then calculated by Microsoft Excel.

5.2.6 Co-immunoprecipitation assays

HEK293T cells were transiently transfected by constructs carrying the DRAG-1 mature region (pCXT239) or the UNC-40 FNIII 5-6 region (pCXT245). Cell media were collected after five days. Anti-c-Myc conjugated sepharose beads (5ul beads per 1.5ml media) were first incubated with the UNC-40 FNIII 5-6 expressing cell media for 4 hours at 4°C, then with the DRAG-1 expressing cell

media for another 4 hours at 4°C. The beads were then subjected to 5 time of wash with 50 mM Tris, pH 8.0, 150 mM NaCl, and 1% Triton X-100, and ran on a 10% of SDS-PAGE gels. The presence of DRAG-1 in the immunoprecipitates was detected by western blotting using mouse anti-FLAG (M2) antibodies (4mg/ml, 1:10000).

5.3 RESULTS

5.3.1 unc-40/neogenin mutants exhibits Sma/Mab pathway mutant phenotypes

We have already shown that *C. elegans* RGM protein DRAG-1 positively modulates the Sma/Mab pathway (Tian et al., 2010). Because neogenin is a receptor for RGM proteins in mouse and in cultured cells (Rajagopalan et al., 2004) and UNC-40 is the sole DCC and neogenin homolog in *C. elegans*, I wished to determine whether UNC-40 plays a role in the Sma/Mab pathway. I first examined *unc-40* mutants for Sma/Mab pathway phenotypes. Previous studies suggested that mutants of the Sma/Mab pathway components exhibit at least three phenotypes: (1) suppression of *sma-9* in patterning the M lineage (Foehr et al., 2006), (2) aberrant body size (Foehr et al., 2006;Savage-Dunn, 2005), and (3) altered expression level of the Sma/Mab responsive reporter RAD-SMAD (Tian et al., 2010). I examined *unc-40* null mutants for all three of these phenotypes.

Two previously published *unc-40* null alleles, *e1430* and *ev457* (Hedgecock et al., 1990;Chan et al., 1996), as well as a previously published null or strong loss-of-function allele *e271* (Brenner, 1974), could suppress the M lineage phenotype of *sma-9(cc604)* mutants at over 70% efficiency (Figure 5.1A, Table 5.1). Furthermore, *unc-40(e1430)* mutant worms have a significantly smaller body size when compared to N2 worms at the same developmental stage (Figure 5.1B). Notably, *unc-40 (e1430)* worms are also significantly smaller than *drag-1(jj4)* worms (Figure 5.1B). Finally, the expression level of the RAD-SMAD reporter is lower in *unc-40(e1430)* background compared to wild type background (Figure 5.2 column 5 from left). It is worth mentioning that RAD-SMAD reporter expression level in *unc-40(e1430)* worms is lower than that in *drag-1(jj4)* worms. This is in consistent with the body size assay, which is also indicative of Sma/Mab signal activity, in which *unc-40(e1430)* worms are smaller than *drag-1(jj4)* worms. Taken together, these results suggest that UNC-40 might function as a positive modulator of the Sma/Mab pathway.

Table 5.1: Mutations in *unc-40* suppress the M lineage defects in the *sma-9* mutants.

Alleles	Molecular lesions	Extent of suppression
unc-40(e1430)	R157STOP (null)	89% (N=150)
unc-40(ev457)	W460STOP (null)	76% (N=106)
unc-40(e271)	R837STOP(null/strong lof)	72% (N=188)
unc-40(ev546)	G774R	12% (N=74)
unc-40(tr115)	W1107STOP	0% (N=166)
unc-40(ur304)	A1056V	0% (N=180)
unc-5(e53)	W283STOP (null)	1.7% (N=59)
unc-6(ev400)	Q78STOP (null)	0.5% (N=221)
unc-6(e78)	C410Y	1% (N=255)

Figure 5.1: *unc-40*, but not *unc-6* or *unc-5*, mutants exhibit a small body size phenotype.

(A) A schematic of UNC-40 protein. It contains four Ig-like and six FNIII domains on the extracellular region and three conserved motifs P1, P2, and P3. The molecular lesions of different mutant alleles are shown in the diagram. (B-C) Body size measurement of different genotypes at the early L4 stage (See materials and methods). *unc-40(e1430)* mutant worms are significantly shorter than wild-type N2, *unc-5(e53), unc-6(ev400),* and *unc-40(tr115)* worms. *unc-40(e1430); lon-1(e185)* and *unc-40(e1430); lon-2(e678)* worms are significantly shorter than *lon-1(e185)* and *lon-2(e678)*, respectively, though to different extent. Error bars represent 95% confidence intervals for the mean body length. The significance of difference between double mutants and the corresponding single mutants was statistically analyzed via Student's t-test. ***, P<0.0001. N.D., not statistically significantly different.

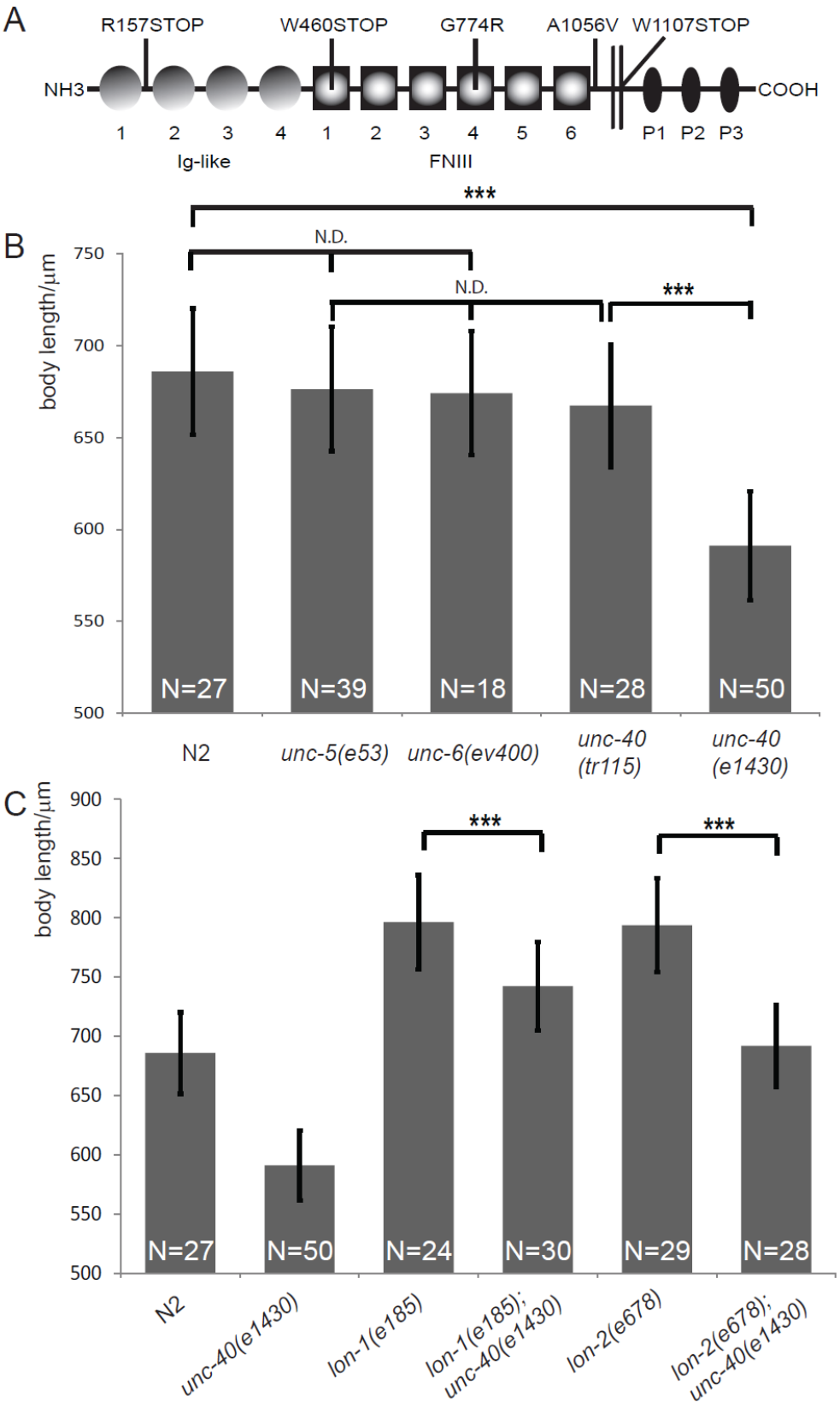
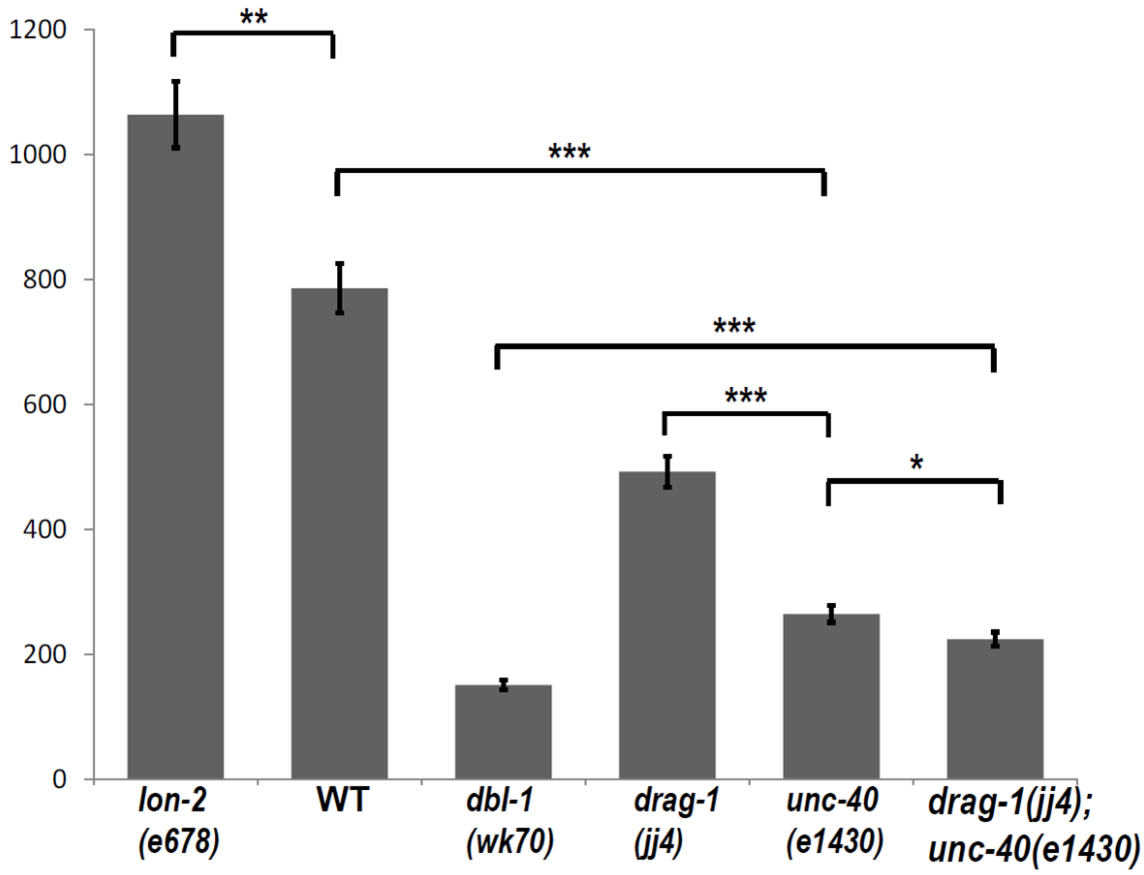

A
R157STOP
W460STOP
G774R
A1056V
W1107STOP
NH3
COOH
1 2 3 4 1 2 3 4 5 6
P1 P2 P3
Ig-like
FNIII
B
***
N.D.
N.D.
***
body length/µm
750
700
650
600
550
500
N=27
N=39
N=18
N=28
N=50
N2
unc-5(e53)
unc-6(ev400)
unc-40 (tr115)
unc-40 (e1430)
C
***
***
body length/µm
900
850
800
750
700
650
600
550
500
N=27
N=50
N=24
N=30
N=29
N=28
N2
unc-40(e1430)
lon-1(e185)
lon-1(e185); unc-40(e1430)
lon-2(e678)
lon-2(e678); unc-40(e1430)

Figure 5.2: Quantification of RAD-SMAD reporter expression in different genetic backgrounds.

Y axis is the signal intensity value. See Materials and Methods for the experimental procedure. Controls include *lon-2(e678)* worms, which have hyperactive Sma/Mab signaling, and *dbl-1(wk70)* worms, which have no Sma/Mab signaling. Error bars represent 95% confidence intervals for the means of relative intensity. The significance of difference among different genotypes was statistically analyzed via Student's t-test. *, $P<0.01$; **, $P<0.001$; ***, $P<0.0001$.


1200
1000
800
600
400
200
0
**
***
***
***
*
lon-2 (e678)
WT
dbl-1 (wk70)
drag-1 (jj4)
unc-40 (e1430)
drag-1(jj4); unc-40(e1430)

5.3.2 unc-6/netrin and unc-5 mutants do not exhibit Sma/Mab pathway mutant phenotypes

Besides RGM, neogenin and DCC are also netrin receptors (Hedgecock et al., 1990;Kennedy et al., 1994;Chan et al., 1996). However, it is not known if netrin or a netrin receptor UNC-5 is involved in BMP signaling. Therefore, I examined the null mutants *unc-6(ev400)* and *unc-5(e53),* for Sma/Mab pathway phenotypes, including the suppression of *sma-9* M lineage defects and body size phenotypes. Unlike *unc-40* null mutants, *unc-6(ev400)* and *unc-5(e53)* null mutants did not suppress the *sma-9* M lineage defects (Table 5.1) and were not significantly different from wild type worms in body length (Figure 5.1B). Together, these results suggest that UNC-40/neogenin may have a role in Sma/Mab signaling that is independent of UNC-6/netrin and UNC-5 in *C. elegans*.

5.3.3 unc-40 genetically interacts with Sma/Mab pathway components

To place UNC-40 in the Sma/Mab pathway, I generated double mutants between *unc-40(e1430)* and mutations in the Sma/Mab pathway and measured their body length. *e678* represents a null allele of *lon-2*, which encodes a member of the glypican family of heparan sulfate proteoglycans and negatively regulates Sma/Mab signaling (Gumienny et al., 2007). *e185* is a strong loss-of-function mutant of *lon-1*, which encodes a downstream target gene of the Sma/Mab pathway that is negatively regulated by the pathway (Maduzia et al., 2002;Morita et al., 2002). However, there is also evidence that *lon-1* may interact with *lon-2* and *dbl-1* via a negative feedback mechanism (King Chow, personal

communication). As shown in Figure 5.1C, *unc-40(e1430); lon-1(e185)* double mutants were significantly smaller than *lon-1(e185)* mutants; *unc-40(e1430); lon-2(e678)* double mutants were also significantly smaller than *lon-2(e678)* mutants. Although *lon-1(e185)* and *lon-2(e678)* mutants are of similar size, *lon-1(e185); unc-40(e1430)* worms are much longer than *lon-2(e678); unc-40(e1430)* worms, suggesting that *unc-40* may function upstream of *lon-1* and in parallel to *lon-2* in the Sma/Mab pathway. A more definitive conclusion requires that additional double mutants, such as *unc-40; sma-3* double, be generated and examined. Since UNC-40 is a transmembrane protein, it likely functions as a modulator of the Sma/Mab pathway at the ligand-receptor level. A possible reason that *lon-1(e185); unc-40(e1430)* double mutants are not as long as *lon-1(e185)* mutants is the proposed negative feedback role of *lon-1* in the Sma/Mab pathway described above.

5.3.4 UNC-40 is expressed in the Sma/Mab signal-receiving cells

To further test the hypothesis that UNC-40 might function as a modulator for the Sma/Mab pathway, we asked whether *unc-40* is expressed in the Sma/Mab signaling-receiving cells. It has been shown that Sma/Mab pathway receptors and Smads function in the hypodermal cells to regulate body size (Yoshida et al., 2001;Wang et al., 2002). We obtained a transgenic line that expresses an integrated functional UNC-40::GFP transgene (Chan et al., 1996). UNC-40::GFP is expressed primarily in neuronal as previously reported (Chan et al., 1996), but is also expressed in the hypodermal cells outlined by anti-AJM-1

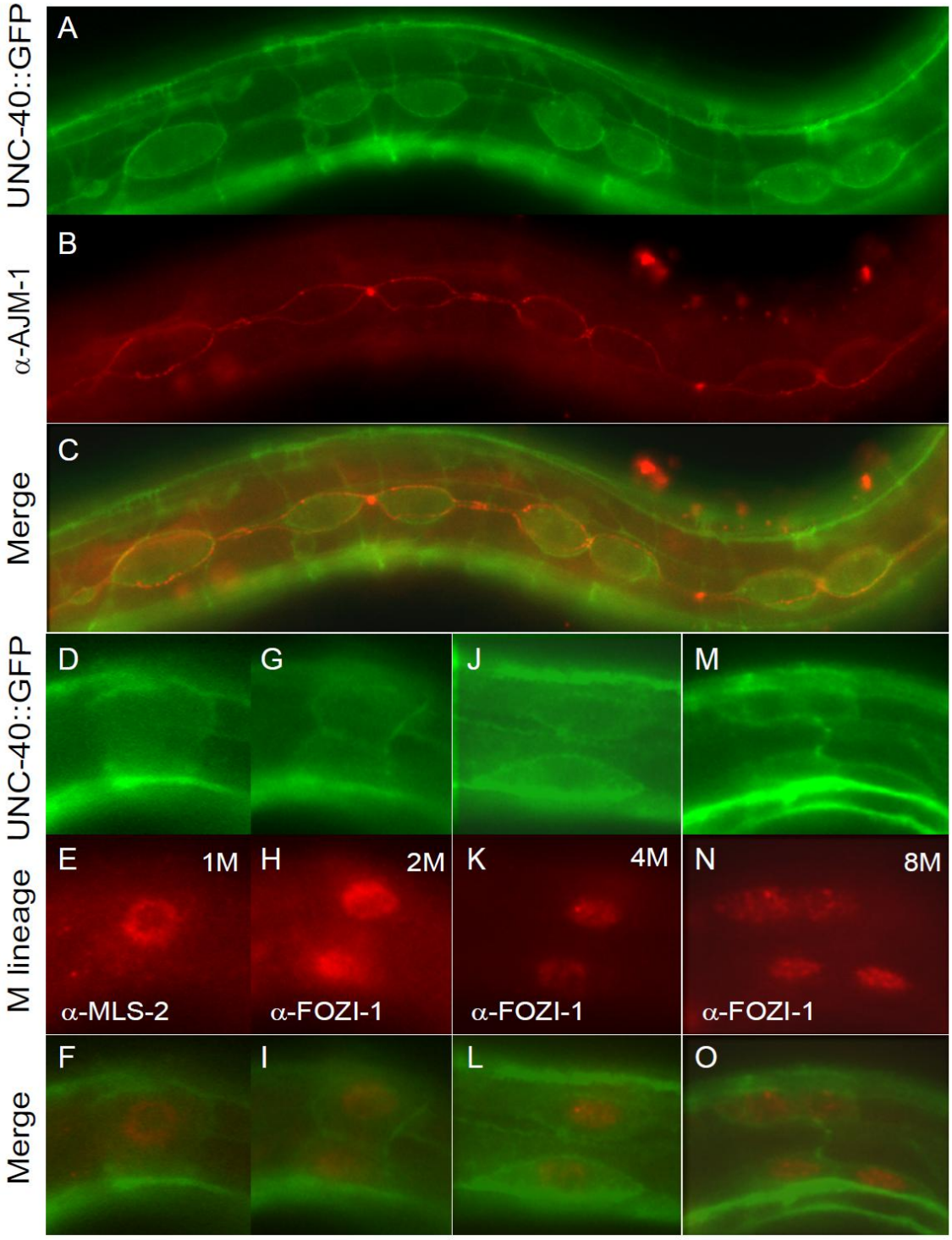
staining (Figure 5.3A-C). Thus, assuming that the transgene faithfully reports expression, UNC-40 is present in the Sma/Mab pathway signal-receiving hypodermal cells.

We also know that the Sma/Mab pathway functions cell-autonomously in the M lineage to control M lineage dorsoventral patterning (Foehr et al., 2006;Tian et al., 2010). Consistent with UNC-40 being a modulator in the Sma/Mab pathway, UNC-40::GFP is expressed in the M lineage at 1-M (Figure 5.3D-F), 2-M (Figure 5.3G-I), 4-M (Figure 5.3J-L), and 8-M (Figure 5.3M-O) stages, which are the critical stages when *sma-9* functions in dorsoventral patterning (Foehr and Liu, 2006). This expression pattern also overlaps with the expression pattern of *drag-1* in the M lineage (Tian et al., 2010).

I am currently testing if forced expression of *unc-40* only in the Sma/Mab signaling-receiving cells can rescue the *unc-40* mutant phenotype in body size regulation and M lineage patterning.

Figure 5.3: *unc-40* is expressed in the hypodermal and M lineage cells

Images shown are side views with anterior to the left and posterior to the right for all panels. UNC-40::GFP (green) is expressed in the hypodermal seam cells (A). Anti-AJM-1 antibody staining (red) labels the apical surface of the hypodermal cells (B). (D-O) The M lineage expression pattern of UNC-40::GFP (green). Anti-MLS-2 antibody staining (red) was used to mark the M mesoblast at the 1-M stage (E). Anti-FOZI-1 antibody staining (red) was used to mark M lineage cells from the 2-M to the 8-M stages (H,K,N). Only one focal plane was shown for the 8-M (N) stage worms. UNC-40::GFP is present from the 1-M to the 8-M stages (D,G,J,M).


UNC-40::GFP
A
α-AJM-1
B
Merge
C
UNC-40::GFP
D
G
J
M
M lineage
E
1M
α-MLS-2
H
2M
α-FOZI-1
K
4M
α-FOZI-1
N
8M
α-FOZI-1
Merge
F
I
L
O

5.3.5 UNC-40 modulates Sma/Mab signaling through its extracellular domain

UNC-40 and its homolog neogenin and DCC have conserved protein domains, including four conserved Immunoglobulin(Ig)-like domains and six Fibronectin type III (FNIII) domains on the extracellular region, a transmembrane domain, and three conserved motifs P1, P2, and P3 intracellularly (Figure 5.1A). The intracellular domain (ICD) of neogenin is known to mediate netrin and RGMa signaling by initiating intracellular signal cascades (Li et al., 2004;Liu et al., 2004;Xie et al., 2005;Ren et al., 2008;Conrad et al., 2007;Hata et al., 2006;Kubo et al., 2008). It is not known whether or not neogenin requires its ICD to mediate BMP signaling. To address this, I tested three hypomorphic *unc-40* alleles, *ev546, ur304* and *tr115*, for their ability to suppress the *sma-9* M lineage defects (Figure 5.1A, Tabel 5.1). Among them, *ev546* and *ur304*, have missense mutations located in the extracellular region (Figure 5.1A). *ev546* is a partial loss-of-function allele in mediating muscle arm extension (Alexander et al., 2010). *ur304* mutant was identified in a suppressor screen and has no defect on its own in cell or axon migration (Xu et al., 2009). I found that *ev546* weakly suppressed, while *ur304* failed to suppress, the *sma-9* M lineage phenotype (Table 5.1). These observations suggest that while the G774R mutation partially compromises the function of UNC-40 in Sma/Mab signaling, the A1056V mutation does not affect the function of UNC-40 in Sma/Mab signaling.

The nonsense mutation (W1107stop) changes the last residue in UNC-40 transmembrane domain into a stop codon, truncating the entire ICD (Figure 5.1A;

Alexander et al., 2009). The mediation of axon and cell movement is dependent on the ICD of UNC-40 (550 Blelloch,R. 1999; 551 Hong,K. 1999), thus *unc-40(tr115)* mutant exhibits an uncoordinated phenotype similar to the *unc-40* null mutants. Surprisingly, *tr115* failed to suppress the *sma-9* M lineage phenotype (Table 5.1), suggesting that the ICD is not required for UNC-40 function in mediating Sma/Mab signaling. Consistent with this result, the body size of *tr115* mutant is not detectably different from that of wild-type, or *unc-5(e53)* and *unc-6(ev400)* null mutants (Figure 5.1B). Taken together, these results indicate that the intracellular domain of UNC-40 is dispensable for its function in Sma/Mab signaling and that the function of UNC-40 in mediating Sma/Mab signaling is independent of its function in axon guidance and cell migration.

5.3.6 UNC-40 interacts with DRAG-1 via its FNIII 5-6 domains and synergizes with DRAG-1 in regulating body size

To understand how UNC-40 functions in modulating Sma/Mab signaling, I examined its biochemical interaction with various components in the Sma/Mab pathway. RGM proteins are known to directly interact with neogenin. This interaction is required not only for multiple roles of RGMa, including axon repulsion (Monnier et al., 2002) and neuronal survival (Koeberle et al., 2010), but also for RGMc/HJV regulating hepcidin expression downstream of BMP signaling (Zhang et al., 2009). Neogenin is known to bind to RGMc/HJV through its FNIII (5,6) domains (Yang et al., 2008;Yang et al., 2011). Therefore, I tested if UNC-40 interacts with the sole *C. elegans* RGM protein DRAG-1 by co-

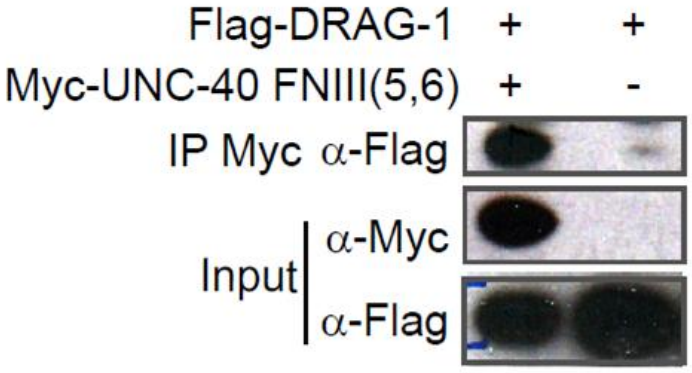
immunoprecipation assays (see materials and methods). Flag-DRAG-1 and Myc-UNC-40 FNIII 5-6 were transfected into mammalian HEK293 cells. IP using anti-Myc antibodies successfully pulled down Flag-DRAG-1 (Figure 5.4A). While I still need to repeat the co-IP experiment in the reciprocal direction and using worm extracts, these results suggest that the DRAG-1 mature domain physically interacts with the FNIII(5,6) domain of UNC-40 protein. I am currently generating truncations and point mutations in UNC-40 to further map the specific residues mediating interaction with DRAG-1.

DRAG-1 is likely a co-receptor of the Sma/Mab pathway (Tian et al., 2010). The physical interaction between DRAG-1 and UNC-40 places both of them in the same complex. Since both of them positively modulate Sma/Mab signaling, I wondered whether they synergize with each other in promoting Sma/Mab signaling. I therefore generated double mutants between *drag-1(jj4)* and *unc-40(e1430)* and determined their body sizes. *drag-1(jj4) unc-40(e1430)* double mutant worms are significantly smaller than either of the single mutant alone (Figure 5.4B). Furthermore, the double mutant has a significantly lower RAD-SMAD reporter activity compared to either of the single mutants (Figure 5.2). Taken together, my biochemical and genetic data suggest that UNC-40 and DRAG-1 likely function in a complex and synergize with each other in modulating Sma/Mab signaling.

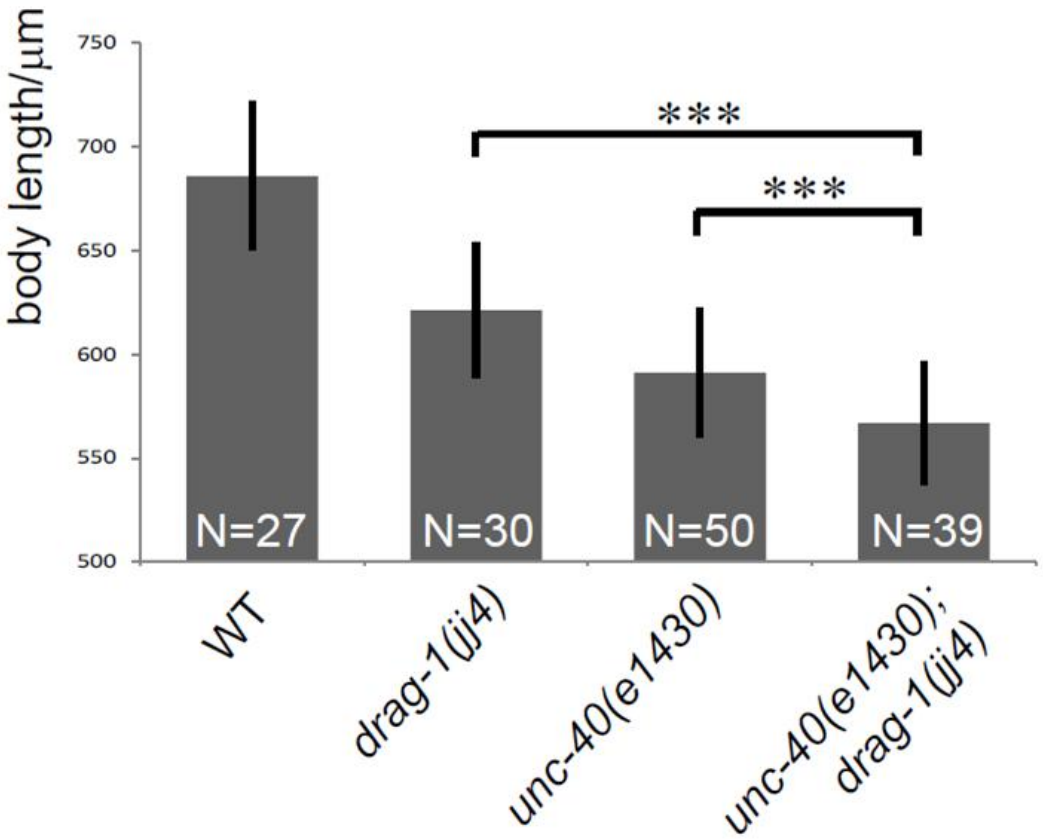
Figure 5.4: UNC-40 physically and genetically interacts with DRAG-1.

(A) UNC-40 co-immunoprecipitates with DRAG-1 via its FNIII(5,6) domain. Please refer to material and methods for experiment procedues. (B) body size measurement of *drag-1(jj4), unc-40(e1430)*, and *drag-1(jj4) unc-40(e1430)* worms. The double mutant is significantly smaller than either single mutant. Error bars represent 95% confidence intervals for the means of relative intensity. The significance of difference between the double mutant and *unc-40(e1430)* was statistically analyzed via Student's t-test. ***, P<0.0001.

A


Flag-DRAG-1 + +
Myc-UNC-40 FNIII(5,6) + -
IP Myc α-Flag
Input
α-Myc
α-Flag

B


body length/μm
750
700
650
600
550
500
***
***
N=27
N=30
N=50
N=39
WT
drag-1(jj4)
unc-40(e1430)
unc-40(e1430); drag-1(jj4)

5.4 DISCUSSION

5.4.1 UNC-40 is a positive modulator of the BMP-like Sma/Mab pathway

As a close homolog of DCC, neogenin was initially recognized as an axon guidance receptor that is required for axon and cell migration (Wilson and Key, 2006;Srinivasan et al., 2003;Kang et al., 2004;Park et al., 2004). Recent *in vitro* and *in vivo* evidence suggests that neogenin is also involved in modulating BMP signaling. Neogenin has been shown to have both positive and negative roles in modulating BMP signaling. Neogenin positively potentiates BMP signaling to regulate the expression of hepcidin for body iron metabolism (Lee et al., 2010;Zhang et al., 2009) and endochondral ossification in mice (Zhou et al., 2010). Neogenin negatively modulates BMP signaling by binding to BMP ligands and activating RhoA signaling in cell cultures (Hagihara et al., 2011). Neogenin also increases the secretion of HJV, which may cause downregulation of BMP signaling (Zhang et al., 2005;Zhang et al., 2007). My data demonstrate that the sole *C. elegans* DCC/neogenin homolog UNC-40 is a positive modulator of the BMP-like Sma/Mab pathway: (1) *unc-40* mutants behave similarly to mutants of all the core members of the Sma/Mab pathway with respect to both body size and suppression of *sma-9* M lineage patterning phenotypes; (2) *unc-40* mutants exhibit reduced activity of the Sma/Mab reporter, RAD-SMAD; (3) genetic epistasis analysis placed *unc-40* at the ligand-receptor level; (4) *unc-40* is expressed in the same cells as the Sma/Mab signal-receiving cells to regulate body size (in hypodermal cells) and mesoderm patterning (in the M lineage); (5)

UNC-40 synergizes with the sole RGM protein DRAG-1 to modulate Sma/Mab signaling and UNC-40 and DRAG-1 likely form a complex by directly interacting with each other. Thus, this study provides *in vivo* evidence that UNC-40/neogenin is a positive modulator of BMP signaling and that this function of neogenin is evolutionarily conserved.

Neogenin is a receptor of RGMa in axon repulsion and the two proteins function in two adjacent cells, or *in trans,* in this process (Rajagopalan et al., 2004). Whether RGM and neogenin function in neighboring cells (*in trans*) or in the same cells (*in cis)* in modulating BMP signaling is also unclear. In this study, I showed that both UNC-40/neogenin and DRAG-1/RGM are expressed in the signal receiving cells, including the hypodermal cells and the M lineage cells. Even though I haven't demonstrated that UNC-40/neogenin functions in these cells, my findings are consistent with UNC-40/neogenin and DRAG-1/RGM functioning in the same cells.

Neogenin is also a receptor for netrin (Wilson and Key, 2006), raising that possibility that netrin or netrin receptors might function in BMP signalling. It hasn't been reported whether netrin or other netrin receptors also participate in BMP signaling. In this study, I showed that *C. elegans* UNC-6/netrin and another netrin receptor UNC-5 do not function in the Sma/Mab pathway: *unc-6* and *unc-5* null mutant have a wild type body size and do not suppress the M lineage defects of *sma-9* mutants. I further showed that UNC-40 protein lacking the ICD exhibits wild-type activity in mediating Sma/Mab signaling. Since this mutant form of

UNC-40 is defective in axon guidance and cell migration mediated by netrin signaling, my results suggest that UNC-40/neogenin exerts its function in mediating BMP signaling via its extracellular domain and that neogenin's role in modulating BMP signaling is independent of its role in netrin signaling. Consistent with this, the netrin-binding motif and the RGM binding motif in neogenin have been mapped to two independent regions of neogenin (Rajagopalan et al. 2004).

5.4.2 A model for how UNC-40 and DRAG-1 function to modulate Sma/Mab signaling

Both UNC-40 and DRAG-1 positively modulate Sma/Mab signaling at the ligand-receptor level (Tian et al., 2010; this study). UNC-40, like its homolog neogenin, appears to directly bind to the RGM protein DRAG-1 through its 5th and 6th FNIII domains. Neogenin-RGM binding has been shown to be required for hepcidin expression by regulating BMP signaling in cell culture (Lee et al., 2010;Zhang et al., 2009). Cell culture experiments have also shown that Neogenin-RGM interaction helps to recruit the BMP receptors to the lipid raft microdomain where signaling occurs (Zhou et al., 2010). Consistent with this notion, the G320V mutation in RGMc/HJV, which disrupts neogenin-HJV interaction, causes juvenile hemochromatosis in humans (Kuns-Hashimoto et al., 2008).

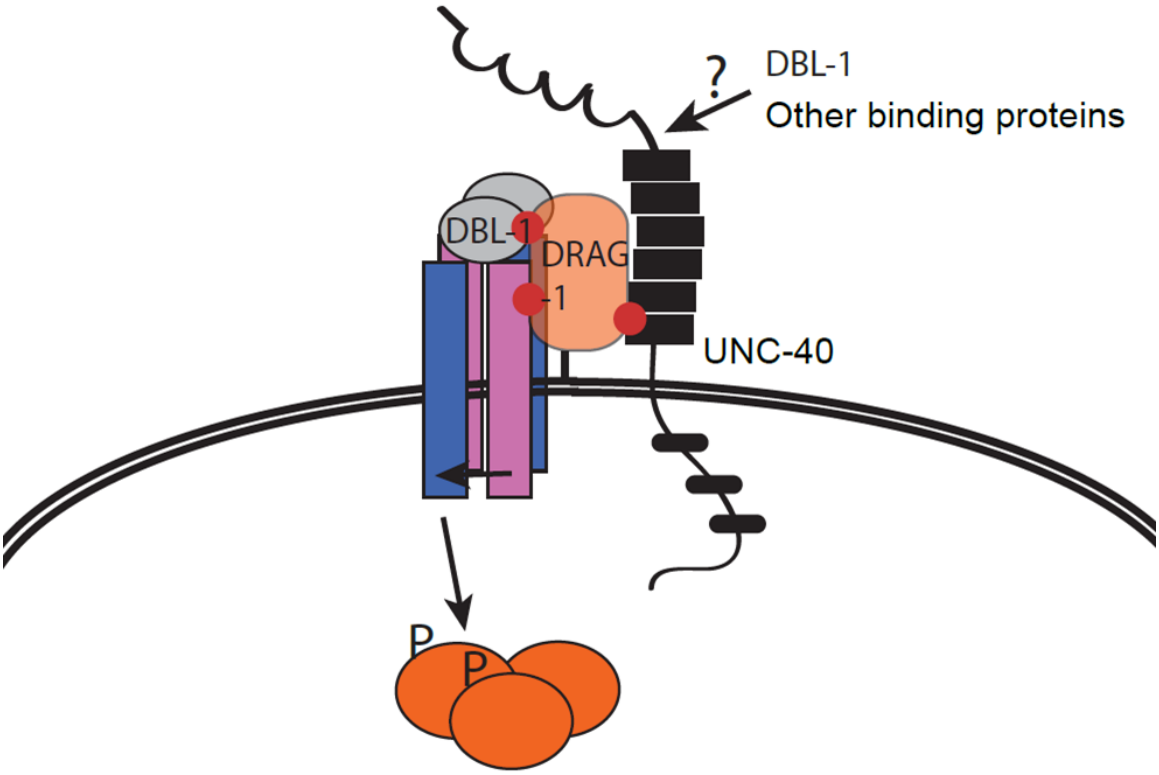
The specific residues in neogenin that are involved in RGM binding haven't been identified. The crystal structure for the 5th and 6th FNIII domains of

neogenin has been solved (Yang et al., 2011). Because neither the 5th FNIII domain in neogenin nor DCC binds RGM (Yang et al., 2008), residues in the 6th FNIII domain of neogenin that are different from those in both the 5th FNIII domain of neogenin and the 6th FNIII domain of DCC are likely to be involved RGM binding. Two loop regions in the 6th FNIII domain of neogenin enriched in these residues are therefore predicted to be involved in neogenin-RGM interaction (Yang et al., 2011). We are now in a position to test the importance of these residues *in vivo*.

It is unlikely that UNC-40/neogenin functions in the BMP-like Sma/Mab pathway by binding to DRAG-1 alone. Recently neogenin has been shown to directly bind to BMP ligands (Hagihara et al., 2011) but not BMP receptors (Zhou et al., 2010). My work provides genetic evidence that *unc-40/neogenin* and *drag-1/RGM* function synergistically, suggesting that *unc-40* may have *drag-1* independent roles. It will be important to find out if UNC-40 also interacts with the ligand and/or receptors of the Sma/Mab pathway. UNC-40/neogenin may be a part of an accessory complex, which at least includes DRAG-1/RGM, which potentiates ligand-receptor binding through local enrichment of the ligand and receptors, which at least includes DRAG-1 (Figure 5.5).

Figure 5.5: A model for how UNC-40/neogenin may function in modulating the BMP-like Sma/Mab pathway

In this model, UNC-40/neogenin positively modulates the BMP-like Sma/Mab pathway through its extracellular domain. The FNIII(5,6) domains of UNC-40/neogenin directly interact with DRAG-1/RGM, which physically interacts with both the DBL-1/BMP ligand and the type I and type II receptors (See Appendix 3). It is unclear if UNC-40/neogenin also physically interacts with the ligand and/or receptors. Similarly, UNC-40/neogenin may also interact with other proteins. Red dots represent directly physical interaction.


?
DBL-1
Other binding proteins
DBL-1
DRAG
-1
UNC-40
P
P

5.5 Acknowledgement

I want to thank Dr. Kang Shen, Dr. Joseph Culotti, Dr. Peter Roy and Dr. Daniel Colon-Ramos for strains and constructs and CGC for strains. I also want to thank Dr. Fenghua Hu for technical help for co-IP experiments. I want to thank Herong Shi for performing molecular cloning in this study.

Chapter 6: Conclusions and Future Perspectives

The proper formation of a complex multicellular organism requires precise coordination among many cellular events, including cell proliferation, cell fate specification and differentiation. I have used the *C. elegans* postembryonic mesodermal lineage, the M lineage, as a tool to study mechanisms coordinating these events at the cellular and molecular resolution, including transcription factor networks and signaling pathways. This thesis has two focuses, uncovering a pro-proliferation regulatory network that is centered with *sem-2/SoxC*, and identifying and characterizing new components in the BMP-like Sma/Mab pathway.

6.1 Identifying a pro-proliferation network in the M lineage (Chapter 2 and Appendix 1)

6.1.1 Summary

In Chapter 2, I described the characterization of the sole SoxC family member in *C. elegans*. I found that SEM-2/SoxC is both necessary and sufficient to promote a proliferating blast cell fate, the sex myoblast fate, over a differentiated striated bodywall muscle fate. This function of *sem-2* is directly regulated by PBC/HOX factors in the M lineage, as demonstrated by computational, *in vivo* functional and *in vitro* electrophoretic mobility shift assays (EMSA). I also identified the positional cues that dictate the specific expression of *sem-2* in the sex myoblast precursors and their descendants, which include SMA-9 antagonism of BMP signaling, Notch signaling and the Wnt/β-catenin asymmetry pathway. The crucial nature of the HOX/PBC factors in directly

enhancing expression of this proliferative factor in the *C. elegans* M lineage suggests a possible more general link between Hox-PBC factors and SoxC proteins in regulating cell proliferation (Tian et al. 2011). In Appendix 1, I described that *sem-4/SALL* functions together with *sem-2/SoxC* in specifying the SMs. Mutations in either *sem-2* or *sem-4* lead to the same SM to BWM fate transformation phenotype, yet they do not regulate each other's expression in the SM mother cells or the SM lineage. SALL proteins are known stem cell factors and regulate differentiation and proliferation in mice. SoxC factors also have roles in proliferation in mice. However, the relationship between SALL and SoxC factors haven't been studied. The result that *sem-2/SoxC* and *sem-4/SALL* together specify the proliferative SMs provides crucial in vivo evidence in any organisms that they may function together.

6.1.2 Future directions

Future directions for understanding how *sem-2* regulates proliferation over differentiation include two parts: understanding the mechanism on how *sem-4* functions together with *sem-2*, and identifying *sem-2* and *sem-4* downstream target genes in the M lineage.

SEM-4 and SEM-2 may physically interact with each other and function as a complex. Yeast two hybrid assays can be used to test for direct interaction. SALL family members are known to interact with NuRD (Nucleosome Remodeling and Deacetylase) and NODE (Nanog- and Oct4-associated Deacetylase) chromatin remodeling complexes (Cox et al., 2010;Kiefer et al., 2002;Liang et al., 2008;van den Berg et al., 2010). Therefore, we can test if the NuRD complex in *C. elegans* is involved in specifying the SMs. To start, we can examine RNAi phenotypes of each component in the NuRD complex. The NuRD

complex may have multiple roles in the M lineage. An early role in general M lineage proliferation and patterning may mask any later roles in specifying the SMs. Therefore, it may be critical to knockdown the NuRD complex components in the M lineage at or before the 8-M stage, when *sem-4* expression is first evident. We can start RNAi treatment at different M lineage stages. We can also use tissue-specific RNAi that employs an RNAi resistant *rde-1* mutant and M lineage-specific rescue with *rde-1* cDNA (Qadota et al., 2007).

SEM-4 and SEM-2 may also regulate the same subset of target genes that are required for the SM fate. To isolate direct target genes of SEM-2 and SEM-4, we can perform Chip-seq on SEM-2 and SEM-4 in isolated M lineage nuclei. The technical hurdle is how to isolate nuclei from *C. elegans* larvae. SM cells from dissociated larvae precociously differentiate, thus the gene expression profiles from such cells wouldn't represent the in vivo situation, which makes cell dissociation not ideal for studying regulatory networks. Our lab is currently working on adapting the INTACT (isolation of nuclei tagged in specific cell types) method (Deal and Henikoff, 2011) to isolate M lineage nuclei, which appears very promising (Vik Ghai, personal communication). In this method, nuclei are biotin-labeled through transgenic expression of a biotinylated nuclear envelope protein in the cell type of interest. Total nuclei are isolated from transgenic worms and biotin-labeled nuclei are then purified using streptavidin-coated magnetic beads. INTACT gives high yield and purity of nuclei from the desired cell types, allowing genome-wide analysis of gene expression and chromatin features. INTACT method has worked for isolating body wall muscle nuclei from *C. elegans* adults (Steiner et al., 2012). If the INTACT method works in the M lineage, we can isolate M lineage nuclei at 8-16M stages in *gfp::sem-2* and *gfp::sem-4* worms. Then we can perform Chip using anti-GFP antibodies on the isolated nuclei and

subject the isolated DNA to high throughput sequencing. We can identify the commonly regulated downstream target genes by both SEM-2 and SEM-4. Gene ontology will be characterized. Those genes whose products belong to protein families that have roles related to proliferation will be subjected to further examination. We will first generate transcriptional and translational reporters to investigate if they are expressed in the M lineage and especially in the proliferating SM cells. The expression will also be examined in *sem-2* and/or *sem-4* mutant background. Loss-of-function and overexpression studiess will then be performed. Those whose expression is regulated by *sem-2* and/or *sem-4* and who have M lineage phenotypes, especially SM proliferation phenotypes are likely the true M lineage specific direct targets of *sem-2* and *sem-4*.

6.2 Identifying and characterizing novel BMP-like Sma/Mab pathway components

6.2.1 Summary

Members of the BMP like signaling pathway in *C. elegans* have been identified from a *sma-9* suppressor screen (Foehr et al. 2006). *drag-1(jj4)* was identified from the same screen. DRAG-1, as the focus of Chapter 3 and Chapter 4, is the sole member of the repulsive guidance molecule (RGM) family of proteins in *C. elegans*. One vertebrate DRAG-1 homolog that is associated with juvenile hemochromatosis (HFE2), Hemojuvilin (Hjv), regulates iron homeostasis by modulating BMP signaling. Using a combination of molecular genetic and biochemical analyses, I demonstrated that DRAG-1 is a membrane-associated protein that functions by interacting with the ligand DBL-1, type I receptor SMA-6, and type II receptor DAF-4 to modulate the Sma/Mab pathway in a cell-type-

specific manner. I further showed that DRAG-1 positively modulates this BMP-like pathway by using a novel Sma/Mab-responsive reporter. I also established that vertebrate Hjv is partially functional in *C. elegans* and three tested HFE2-associated mutations hamper DRAG-1 function *in vivo*. Therefore, my results provide a direct link between RGM proteins and BMP signaling *in vivo* and a simple and genetically tractable system for mechanistic studies of RGM protein regulation of BMP pathways.

In Chapter 5, I described the role of UNC-40, the sole Neogenin/DCC protein, in the Sma/Mab pathway. This role of neogenin has been shown in a couple of BMP-mediated processes in mice and cultured mammalian cells; however, mechanistically how neogenin functions in BMP signaling is still not clear. I genetically placed *unc-40* in the BMP-like Sma/Mab pathway. I showed that UNC-40 positively regulates Sma/Mab signaling with the RAD-SMAD reporter. Furthermore, my results demonstrated that modulation of Sma/Mab signaling by UNC-40 is independent of its role in axon and cell migration: mutations in UNC-6/netrin and UNC-5, factors that are related to the axon guidance role of UNC-40, do not exhibit the Sma/Mab pathway phenotypes. Interestingly, I found that the ectodomain of UNC-40 is sufficient to mediate its role in BMP signaling because a mutant allele of *unc-40* carrying an early stop codon that truncates the entire intracellular region could function perfectly in BMP signaling but not in axon guidance. UNC-40 physically interacts with DRAG-1 through UNC-40's FNIII (5,6) domains in the extracellular region, which suggests that the UNC-40-DRAG-1 complex may be required for modulating Sma/Mab signaling. Consistent with this notion, I showed that one HFE2-associated mutation in Hjv, known to interrupt HJV binding to neogenin, almost completely abolished DRAG-1 function in *C. elegans*.

6.2.2 Future directions

The main future focus will be investigating the mechanisms on how UNC-40 and DRAG-1 modulate Sma/Mab signaling.

One hypothesis is that UNC-40 and DRAG-1 interaction is required for their function in Sma/Mab signaling. We can identify the critical residues in both UNC-40 and DRAG-1 proteins that are mediating the interaction. The Neogenin-interacting region in RGMa protein has been mapped (Itokazu et al., 2012). Synthesized peptide consisting of aa 284-293 of RGMa directly binds to the extracellular domain (ECD) of recombinant neogenin. The homologous sequence in DRAG-1 is aa238-247, among which 5 out of the 10 residues are identical to RGMa (Figure 4.1, blue underline). Therefore, this stretch of amino acids provides us with a starting point to test critical residues in DRAG-1 that are required for DRAG-1-UNC-40 binding. Missense mutations will be made in DRAG-1 protein in the critical residues. Mutated DRAG-1 will be tested for function regarding body size and *sma-9* M lineage suppression phenotypes. One human JH mutation, G320V, is known to interrupt HJV binding to neogenin (ref). G320 corresponds to G271 in DRAG-1. I have already shown that the G271V mutation in DRAG-1 almost completely abolished DRAG-1 function in mediating Sma/Mab signaling. Notably, G271 is located outside of the 10-residue motif (aa238-247). We will test if ^{G271V}DRAG-1 is able to bind to UNC-40.

I already showed that UNC-40 binds to DRAG-1 through its FNIII(5,6) domains. The crystal structure of neogenin FNIII(5,6) domains has been solved (Yang et al., 2011a). Because neither the neogenin 5th FNIII domain nor DCC directly binds RGM (Yang et al., 2008), residues in the neogenin 6th FNIII domain that are different from residues in both the neogenin 5th FNIII domain and the DCC 6th FNIII domain are likely the ones required for RGM binding. Two loops

enriched with the differential residues were therefore predicted to mediate neogenin-RGM interaction (Yang et al., 2011b). Since UNC-40 is a homolog for both DCC and neogenin, the residues in UNC-40 that are conserved with neogenin but not DCC in those loops are most likely to be candidates for neogenin-RGM interaction. These candidate residues could be subjected to combined biochemical interaction and in vivo function tests to investigate if DRAG-1-UNC-40 interaction is important for UNC-40 function.

It is likely that UNC-40 and/or DRAG-1 mediate Sma/Mab signaling not just through DRAG-1-UNC-40 interaction, because *unc-40(e1430); drag-1(jj4)* double mutants have more severe Sma/Mab pathway phenotypes. Neogenin has been shown to directly interact with the BMP ligands (Hagihara et al., 2011). Therefore, we would like to test if UNC-40 also binds to the ligand DBL-1, if so, we will determine which part of UNC-40 is required for such interaction. We would also like to study the functional relevance of this interaction.

If DRAG-1-UNC-40 interaction is required for them to mediate Sma/Mab signaling, the next question will be how they promote the signaling. DRAG-1 can physically interact with the ligand and both type I and type II receptors in the Sma/Mab pathway (Chapter 4). The hypothesis is that the formation of a supercomplex comprised of the ligand, type I, II receptors, DRAG-1, and UNC-40, will cause local enrichment of both the ligand and receptors on the membrane, which can lead to efficient signaling. One recent study in cultured cells suggested that neogenin can recruit the receptors to the lipid rafts microdomain where the signaling takes place, with the help of RGM proteins (Zhou et al., 2010). We would like to test if UNC-40 and DRAG-1 also help the receptors enter the lipid rafts in *C. elegans*. We will first examine the presence of receptors, DRAG-1 and UNC-40 in lipid rafts in wild type animals by isolating lipid

rafts from transgenic animals that express gfp fusions for each protein. If they are present in lipid rafts in wild type animals, we will then isolate lipid rafts from *drag-1(jj4)*, and *unc-40(e1430)* mutants and examine the presence of receptors in lipid rafts.

The Liu lab has performed a large scale *sma-9* suppressor screen (Jun, Liu, unpublished results). We have identified many mutations that are not located in the known Sma/Mab pathway genes. The long term goal will be identifying and characterizing the new components using methods similar to those I used in Chapters 3-5. The new components, especially ones that are localized to the cellular membrane, will add to our current understanding of RGM-neogenin complex in modulating BMP signaling.

APPENDIX I:

Determine the relationship between *sem-2/SoxC* and *sem-4/SALL* in the M lineage

A1.1 INTRODUCTION

sem-2 encodes the sole *C. elegans* group C Sry-related HMG box (Sox)-containing transcription factor . Previous study has shown that SEM-2 is necessary to promote a proliferating sex myoblast (SM) fate over a differentiated striated body wall muscle (BWM) fate (Tian et al., 2011). *sem-2* expression in the sex myoblast precursors and their descendants is under the control of a number of factors, including direct activation by a Hox-PBC complex, regulation by the Wnt/β-catenin asymmetry pathway along the anterioposterior axis, as well as regulation by the dorsoventral cues: SMA-9 antagonism of the Sma/Mab signaling that patterns the dorsal M lineage and Notch signaling that patterns the ventral M lineage, both of which converge on a FoxF/C factor LET-381 that represses *sem-2* expression. The function of SEM-2 in the M lineage suggests that SoxC proteins have an evolutionarily conserved role in promoting the proliferation of multipotent progenitor cells. It is crucial for our understanding of SoxC function to identify the co-factors and downstream target genes through which SoxC proteins exert their function in regulating progenitor cell proliferation.

It has been shown that *sem-4* mutants share the same SM to BWM fate transformation phenotype as *sem-2* mutants; in addition, *sem-4* mutants exhibit an extra division in the coelomocyte (CC) precursor cells, which then give rise to two BWMs (Basson and Horvitz, 1996). SEM-4 is the sole member of the Spalt or SALL family of C2H2 zinc finger-containing transcription factors in *C. elegans*

(de Celis and Barrio, 2009). The SALL family contains four members in vertebrates, SALL1-SALL4. SALL family members have been shown to play roles in regulating proliferation in vertebrates. Mutations in SALL1 and SALL4 in humans cause Townes-Brocks syndrome and Okihiro syndrome, respectively, where patients exhibit hypoplastic phenotypes in multiple tissues/organs (de Celis and Barrio, 2009;Sweetman and Munsterberg, 2006). Both SALL1 and SALL4 have been shown to regulate pluripotency of stem or progenitor cells (Karantzali et al., 2011;Neff et al., 2011;Oikawa et al., 2009;Rao et al., 2010;Wang et al., 2011;Yang et al., 2008;Yang et al., 2010;Yang et al., 2012;Yuri et al., 2009). Mechanistically, SALL proteins function by interacting with NuRD (Nucleosome Remodeling and Deacetylase) and NODE (Nanog- and Oct4-associated Deacetylase) chromatin remodeling complexes (Cox et al., 2010;Kiefer et al., 2002;Liang et al., 2008;van den Berg et al., 2010). SoxC proteins share similar roles in the self-renewal of multipotent progenitor cells (See Chapter 1). However, the functional link between SoxC and SALL proteins in regulating proliferation is unknown. The shared M lineage phenotype between *sem-2/SoxC* and *sem-4/SALL* mutants suggests an exciting possibility that the functions of these two factors are linked.

A1.2 MATERIAL AND METHODS

A1.2.1 C. elegans strains

The following strains were used in this study: LG I: *sem-2(n1343)*; *sem-4(n1378)*. Analyses were performed at 20°C.

A1.2.2 Immunofluorescence staining

Animal fixation, immunostaining, microscopy and image analysis were performed as described previously (Amin et al., 2007). Guinea pig anti-FOZI-1 (Amin et al., 2007) (1:200) and goat anti-GFP (Rockland Immunochemicals; 1:1000) were used.

A1.3 RESULTS

A1.3.1 sem-4 shares overlapping expression pattern with sem-2 in the M lineage

I examined *sem-4* expression in the M lineage using a *sem-4::gfp* transgene (Grant et al., 2000). On the ventral side, *sem-4::gfp* expression was first detected at the 8-M stage in the M.v(l/r)p cells, or the SM grandmother cells (50%, N=6). This expression persisted in the SM mother cells (61.5%, N=39), the sister cell of the SM mother cell (56.4%, N=39), and the SMs (100%, N=8, Figure A1.2A,B). The expression in SM descendants hasn't been investigated. The differentiating BWMs and CCs gradually lose *sem-4::gfp* expression. Thus *sem-4::gfp* expression precedes that of *sem-2::gfp* expression by one cell division in the ventral M lineage, where they subsequent share overlapping expression patterns (Figure A1.2A,B., Tian et al., 2011). On the dorsal side, *sem-4::gfp* expression was first observed in the M.d(l/r)p cells, or the CC mother cells (60%, N=10) and subsequently in their daughter cells, or the undifferentiated CCs (54.5%, N=11) and BWMs (45,5%, N=11, Figure A1.2B). *sem-4::gfp* expression

was not detectable in the differentiating CCs and BWMs. Interestingly, this loss of *sem-4::gfp* expression in the differentiating cells appeared to be transient, as *sem-4::gfp* expression came back on in the fully differentiated BWMs and CCs (data not shown). The functional significance of this later expression of *sem-4::gfp* is not known.

Figure A1. 1: *sem-4(n1378)* M lineage phenotypes.

(A-B) On the dorsal side, *sem-4* mutant has the CC precusors going through an extra division and give rise to a pair of BWMs. On the ventral side, *sem-4* mutant has a fate transformation from SMs to BWMs.

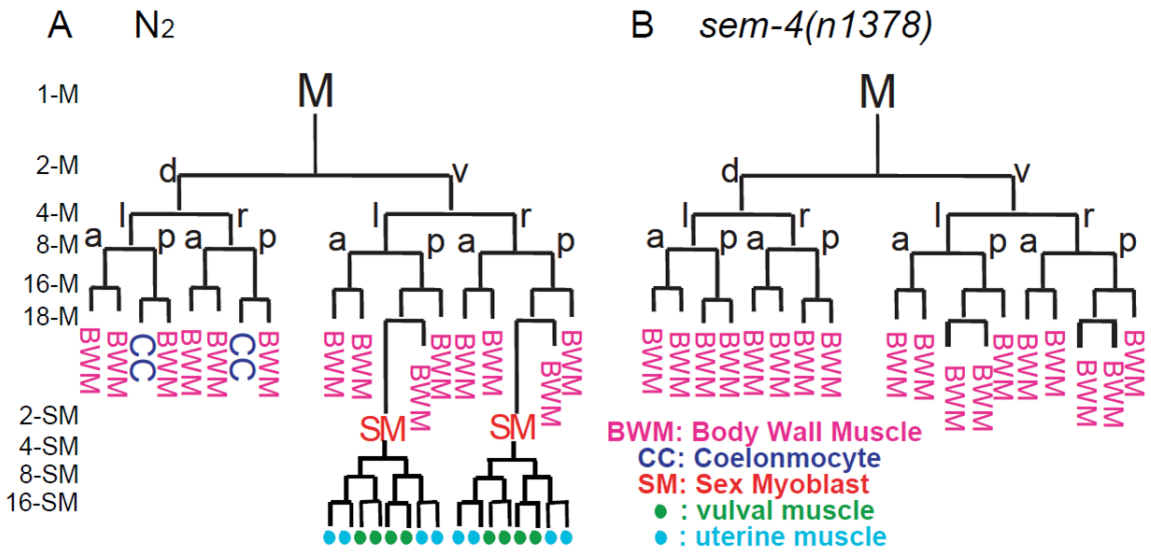
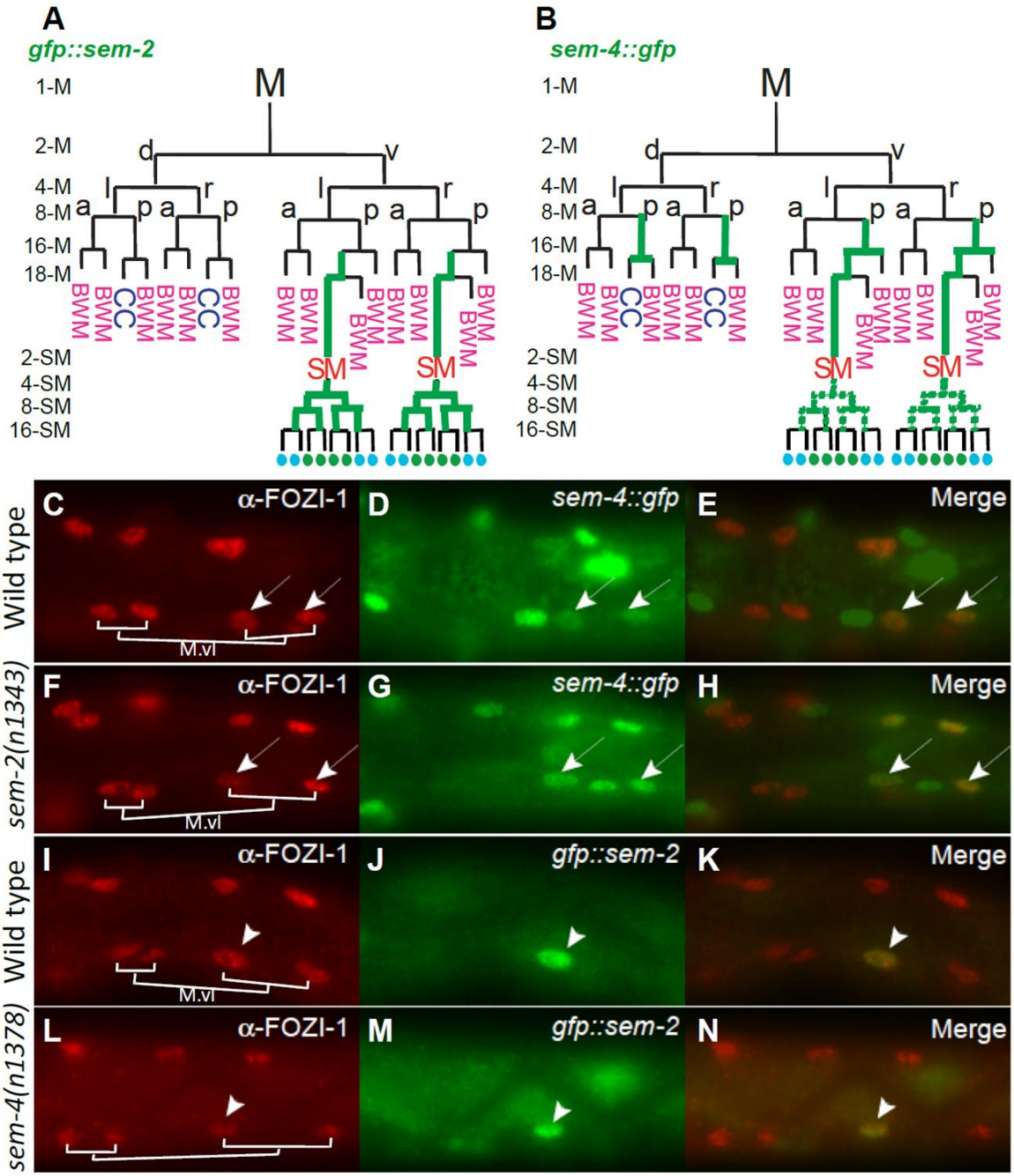

A N2
B sem-4(n1378)
1-M
2-M
4-M
8-M
16-M
18-M
2-SM
4-SM
8-SM
16-SM
M
d
v
l
r
a
p
BWM
CC
SM
BWM: Body Wall Muscle
CC: Coelonmocyte
SM: Sex Myoblast
: vulval muscle
: uterine muscle

Figure A1. 2: In the M lineage, *sem-2* and *sem-4* share over lapping expression patterns, but they do not regulate each other's expression.

The expression patterns of *sem-4::gfp* and *gfp::sem-2* in the M lineage are summarized in panels A and B respectively. On the ventral side, *gfp::sem-2* is expressed in the SM mother cells and the SM lineage; *sem-4::gfp* expression is detected a cell division earlier than that of *sem-2* and overlaps with *sem-2* in the SM mother and SM cells. *sem-4::gfp* expression hasn't been examined in the SM descendants, as shown in the dotted lines in panel B. *sem-4::gfp* is also expressed on the dorsal side in CC mother cells, CC precursors and their sister BWM precursor cells, while *gfp::sem-2* is not expressed on the dorsal side. M lineage cells are labeled with anti-FOZI-1 antibody (C, F, I, L). Expression patterns of *sem-4::gfp* are similar in wild-type and *sem-2(n1343)* worms (C-H), while expression patterns of *gfp::sem-2* are similar between wild-type and *sem-4(n1378)* worms (I-N). Arrows point to the GFP positive M lineage cells.


A
gfp::sem-2
B
sem-4::gfp
1-M
2-M
4-M
8-M
16-M
18-M
2-SM
4-SM
8-SM
16-SM
M
d
v
l
r
a
p
BWM
CC
SM
C
α-FOZI-1
D
sem-4::gfp
E
Merge
Wild type
M.vl
F
α-FOZI-1
G
sem-4::gfp
H
Merge
sem-2(n1343)
M.vl
I
α-FOZI-1
J
gfp::sem-2
K
Merge
Wild type
M.vl
L
α-FOZI-1
M
gfp::sem-2
N
Merge
sem-4(n1378)

A1.3.2 sem-4 and sem-2 do not regulate each other's expression in the ventral M lineage

Because *sem-4* shares the same ventral M lineage phenotype and expression with *sem-2*, I determined whether they regulate each other's expression in the ventral M lineage. I used the *sem-4(n1378)* allele, which showed 98% SM to BWM fate transformation (Basson et al., 1996), and the *sem-2(n1343)* allele, which showed 100% SM-to-BWM fate transformation (Tian et al., 2011). To my surprise, the expression of *gfp::sem-2* was still detectable in *sem-4(n1378)* mutants in the M lineage at the 16-M stage (93.75%, N=32) and the 18-M stage (100%, N=6), in the same pattern and intensity as compared to wild type (Figure A1.2 I-N). Similarly, the expression of *sem-4::gfp* was also present in *sem-2(n1343)* mutants in the M lineage at the 16-M stage (100%, N=8) and the 18-M stage (100%, N=3), in the same pattern and intensity as compared to wild type (Figure A1.2 C-H). These results suggest that *sem-2* and *sem-4* do not regulate each other's expression.

sem-2::gfp is not expressed in the dorsal M lineage in wild type worms. Interesting, it is expressed in the M.d(l/r)pa cells in *sem-4(n1378)* mutants (34.7%, N=23). *sem-4(n1378)* mutant worms exhibit a 30% dorsal CC-to -BWM fate transformation phenotype (Basson et al., 1996). Taken together, these observations suggest that *sem-2* is normally repressed by *sem-4* in the dorsal M lineage.

A1.4 DISCUSSION

Both *sem-2* and *sem-4* are expressed in the proliferating SM precursors and descendants (Tian et al., 2011; this work), while *sem-2* and *sem-4* mutants fail to make the proliferating SM cells (Tian et al., 2011; Basson et al., 1996). I have shown that these two genes do not regulate each other's expression in the ventral M lineage. These findings suggest that SEM-2 and SEM-4 are required together to specify the proliferative SM fate. It is important to determine how they function together. One possibility is that SEM-2 and SEM-4 directly physically interact with each other in regulating gene expression; another possibility is that they regulate the same set of target genes involved in SM specification and proliferation. Despite the lack of mechanistic details, these studies begin to connect the SoxC factors with the SALL stem cell factors. Since SALL protein function involves the NuRD complex (Cox et al., 2010;Kiefer et al., 2002;Liang et al., 2008;van den Berg et al., 2010), it also be interesting to investigate if the *C. elegans* NuRD complex is involved in SEM-4 and SEM-2 function in SM specification.

In addition to the SM-to-BWM phenotype, *sem-4* mutants also share similar but not identical phenotypes as *let-381(RNAi)* animals in the dorsal M lineage. *let-381* encodes the sole FoxF/C factor in *C. elegans*, and in *let-381(RNAi)* worms the M.d(l/r)pa cells (which differentiate into CCs in wild-type worms) behave as their ventral counterparts, the M.v(l/r)pa cells that divide to produce a SM and a BWM each (Amin et al., 2010). In *sem-4* mutants, the M.d(l/r)pa cells divide an extra time like their ventral counterparts (a phenotype

shared by *let-381(RNAi)* and *sem-4* mutants), but give rise to 4 BWMs, rather than 2 SMs and 2 BWMs like in *let-381(RNAi)* animals (Figure A1.1). It is known that *let-381* represses *sem-2* expression on the dorsal M lineage (Tian et al., 2011). *sem-4* also appears to repress *sem-2* in the dorsal M lineage (data not shown). Thus, the extra division phenotype of the M.d(l/r)pa cells in both *let-381(RNAi)* and *sem-4(n1378)* animals might be due to the ectopic expression of *sem-2* in these cells. The reason that the anterior daughters of these cells become the SMs in *let-381(RNAi)* animals is probably because both *sem-2* and *sem-4* are present in these cells. However, the lack of *sem-4* in *sem-4(n1378)* mutants will render these cells to become BWMs because both *sem-2* and *sem-4* are required for the proper specification of SMs. Therefore, it is important to study the relationship between *sem-4* and *let-381* in the dorsal M lineage. One prediction is that *sem-4* either positively regulates *let-381* expression or works in parallel to *let-381*, while *let-381* does not regulate the expression of *sem-4*.

References

Aaboe, M., Birkenkamp-Demtroder, K., Wiuf, C., Sorensen, F. B., Thykjaer, T., Sauter, G., Jensen, K. M., Dyrskjot, L. and Orntoft, T. (2006). SOX4 Expression in Bladder Carcinoma: Clinical Aspects and *in vitro* Functional Characterization. *Cancer Res.* 66, 3434-3442.

Alexander, M., Chan, K. K., Byrne, A. B., Selman, G., Lee, T., Ono, J., Wong, E., Puckrin, R., Dixon, S. J. and Roy, P. J. (2009). An UNC-40 Pathway Directs

Postsynaptic Membrane Extension in Caenorhabditis Elegans. *Development* 136, 911-922.

Alexander, M., Selman, G., Seetharaman, A., Chan, K. K., D'Souza, S. A., Byrne, A. B. and Roy, P. J. (2010). MADD-2, a Homolog of the Opitz Syndrome Protein MID1, Regulates Guidance to the Midline through UNC-40 in Caenorhabditis Elegans. *Dev. Cell.* 18, 961-972.

Ambrosio, A. L., Taelman, V. F., Lee, H. X., Metzinger, C. A., Coffinier, C. and De Robertis, E. M. (2008). Crossveinless-2 is a BMP Feedback Inhibitor that Binds Chordin/BMP to Regulate Xenopus Embryonic Patterning. *Dev. Cell.* 15, 248-260.

Amin, N. M., Hu, K., Pruyne, D., Terzic, D., Bretscher, A. and Liu, J. (2007). A Zn-finger/FH2-Domain Containing Protein, FOZI-1, Acts Redundantly with CeMyoD to Specify Striated Body Wall Muscle Fates in the Caenorhabditis Elegans Postembryonic Mesoderm. *Development* 134, 19-29.

Amin, N. M., Lim, S. E., Shi, H., Chan, T. L. and Liu, J. (2009). A Conserved Six-Eya Cassette Acts Downstream of Wnt Signaling to Direct Non-Myogenic Versus Myogenic Fates in the *C. elegans* Postembryonic Mesoderm. *Dev. Biol.* 331, 350-360.

Amin, N. M., Shi, H. and Liu, J. (2010). The FoxF/FoxC Factor LET-381 Directly Regulates both Cell Fate Specification and Cell Differentiation in *C. elegans* Mesoderm Development. *Development* 137, 1451-1460.

Andersen, C. L., Christensen, L. L., Thorsen, K., Schepeler, T., Sorensen, F. B., Verspaget, H. W., Simon, R., Kruhoffer, M., Aaltonen, L. A., Laurberg, S. et al. (2009). Dysregulation of the Transcription Factors SOX4, CBFB and SMARCC1 Correlates with Outcome of Colorectal Cancer. *Br. J. Cancer* 100, 511-523.

Andriopoulos, B.,Jr, Corradini, E., Xia, Y., Faasse, S. A., Chen, S., Grgurevic, L., Knutson, M. D., Pietrangelo, A., Vukicevic, S., Lin, H. Y. et al. (2009). BMP6 is a Key Endogenous Regulator of Hepcidin Expression and Iron Metabolism. *Nat. Genet.* 41, 482-487.

Aragon, E., Goerner, N., Zaromytidou, A. I., Xi, Q., Escobedo, A., Massague, J. and Macias, M. J. (2011). A Smad Action Turnover Switch Operated by WW Domain Readers of a Phosphoserine Code. *Genes Dev.* 25, 1275-1288.

Arata, Y., Kouike, H., Zhang, Y., Herman, M. A., Okano, H. and Sawa, H. (2006). Wnt Signaling and a Hox Protein Cooperatively Regulate Psa-3/Meis to Determine Daughter Cell Fate After Asymmetric Cell Division in *C. elegans*. *Dev. Cell.* 11, 105-115.

Argiropoulos, B. and Humphries, R. K. (2007). Hox Genes in Hematopoiesis and Leukemogenesis. *Oncogene* 26, 6766-6776.

Attisano, L. and Wrana, J. L. (2000). Smads as Transcriptional Co-Modulators. *Curr. Opin. Cell Biol.* 12, 235-243.

Attisano, L., Wrana, J. L., Lopez-Casillas, F. and Massague, J. (1994). TGF-Beta Receptors and Actions. *Biochim. Biophys. Acta* 1222, 71-80.

Babitt, J. L., Huang, F. W., Wrighting, D. M., Xia, Y., Sidis, Y., Samad, T. A., Campagna, J. A., Chung, R. T., Schneyer, A. L., Woolf, C. J. et al. (2006). Bone Morphogenetic Protein Signaling by Hemojuvelin Regulates Hepcidin Expression. *Nat. Genet.* 38, 531-539.

Babitt, J. L., Huang, F. W., Xia, Y., Sidis, Y., Andrews, N. C. and Lin, H. Y. (2007). Modulation of Bone Morphogenetic Protein Signaling *in vivo* Regulates Systemic Iron Balance. *J. Clin. Invest.* 117, 1933-1939.

Babitt, J. L., Zhang, Y., Samad, T. A., Xia, Y., Tang, J., Campagna, J. A., Schneyer, A. L., Woolf, C. J. and Lin, H. Y. (2005). Repulsive Guidance Molecule (RGMa), a DRAGON Homologue, is a Bone Morphogenetic Protein Co-Receptor. *J. Biol. Chem.* 280, 29820-29827.

Bae, G. U., Yang, Y. J., Jiang, G., Hong, M., Lee, H. J., Tessier-Lavigne, M., Kang, J. S. and Krauss, R. S. (2009). Neogenin Regulates Skeletal Myofiber Size and Focal Adhesion Kinase and Extracellular Signal-Regulated Kinase Activities *in vivo* and *in vitro*. *Mol. Biol. Cell* 20, 4920-4931.

Balemans, W. and Van Hul, W. (2002). Extracellular Regulation of BMP Signaling in Vertebrates: A Cocktail of Modulators. *Dev. Biol.* 250, 231-250.

Basson, M. and Horvitz, H. R. (1996). The Caenorhabditis Elegans Gene Sem-4 Controls Neuronal and Mesodermal Cell Development and Encodes a Zinc Finger Protein. *Genes Dev.* 10, 1953-1965.

Bergsland, M., Ramskold, D., Zaouter, C., Klum, S., Sandberg, R. and Muhr, J. (2011). Sequentially Acting Sox Transcription Factors in Neural Lineage Development. *Genes Dev.* 25, 2453-2464.

Bergsland, M., Werme, M., Malewicz, M., Perlmann, T. and Muhr, J. (2006). The Establishment of Neuronal Properties is Controlled by Sox4 and Sox11. *Genes Dev.* 20, 3475-3486.

Bhattaram, P., Penzo-Mendez, A., Sock, E., Colmenares, C., Kaneko, K. J., Vassilev, A., Depamphilis, M. L., Wegner, M. and Lefebvre, V. (2010). Organogenesis Relies on SoxC Transcription Factors for the Survival of Neural and Mesenchymal Progenitors. *Nat. Commun.* 1, 9.

Blelloch, R., Newman, C. and Kimble, J. (1999). Control of Cell Migration during Caenorhabditis Elegans Development. *Curr. Opin. Cell Biol.* 11, 608-613.

Bohme, U. and Cross, G. A. (2002). Mutational Analysis of the Variant Surface Glycoprotein GPI-Anchor Signal Sequence in Trypanosoma Brucei. *J. Cell. Sci.* 115, 805-816.

Bowles, J., Schepers, G. and Koopman, P. (2000). Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators. *Dev. Biol.* 227, 239-255.

Brennan, D. J., Ek, S., Doyle, E., Drew, T., Foley, M., Flannelly, G., O'Connor, D. P., Gallagher, W. M., Kilpinen, S., Kallioniemi, O. P. et al. (2009). The Transcription Factor Sox11 is a Prognostic Factor for Improved Recurrence-Free Survival in Epithelial Ovarian Cancer. *Eur. J. Cancer* 45, 1510-1517.

Brenner, S. (1974). The Genetics of Caenorhabditis Elegans. *Genetics* 77, 71-94.

Broitman-Maduro, G., Maduro, M. F. and Rothman, J. H. (2005). The Noncanonical Binding Site of the MED-1 GATA Factor Defines Differentially Regulated Target Genes in the *C. elegans* Mesendoderm. *Dev. Cell.* 8, 427-433.

Camaschella, C., Roetto, A. and De Gobbi, M. (2002). Juvenile Hemochromatosis. *Semin. Hematol.* 39, 242-248.

Camus, L. M. and Lambert, L. A. (2007). Molecular Evolution of Hemojuvelin and the Repulsive Guidance Molecule Family. *J. Mol. Evol.* 65, 68-81.

Chan, S. K. and Mann, R. S. (1996). A Structural Model for a Homeotic Protein-Extradenticle-DNA Complex Accounts for the Choice of HOX Protein in the Heterodimer. *Proc. Natl. Acad. Sci. U. S. A.* 93, 5223-5228.

Chan, S. S., Zheng, H., Su, M. W., Wilk, R., Killeen, M. T., Hedgecock, E. M. and Culotti, J. G. (1996). UNC-40, a *C. elegans* Homolog of DCC (Deleted in Colorectal Cancer), is Required in Motile Cells Responding to UNC-6 Netrin Cues. *Cell* 87, 187-195.

Chatterjee, S. and Mayor, S. (2001). The GPI-Anchor and Protein Sorting. *Cell Mol. Life Sci.* 58, 1969-1987.

Colavita, A. and Culotti, J. G. (1998). Suppressors of Ectopic UNC-5 Growth Cone Steering Identify Eight Genes Involved in Axon Guidance in Caenorhabditis Elegans. *Dev. Biol.* 194, 72-85.

Colavita, A., Krishna, S., Zheng, H., Padgett, R. W. and Culotti, J. G. (1998). Pioneer Axon Guidance by UNC-129, a *C. elegans* TGF-Beta. *Science* 281, 706-709.

Cole, S. J., Bradford, D. and Cooper, H. M. (2007). Neogenin: A Multi-Functional Receptor Regulating Diverse Developmental Processes. *Int. J. Biochem. Cell Biol.* 39, 1569-1575.

Conrad, S., Genth, H., Hofmann, F., Just, I. and Skutella, T. (2007). Neogenin-RGMa Signaling at the Growth Cone is Bone Morphogenetic Protein-Independent and Involves RhoA, ROCK, and PKC. *J. Biol. Chem.* 282, 16423-16433.

Conrad, S., Stimpfle, F., Montazeri, S., Oldekamp, J., Seid, K., Alvarez-Bolado, G. and Skutella, T. (2010). RGMb Controls Aggregation and Migration of Neogenin-Positive Cells *in vitro* and *in vivo*. *Mol. Cell. Neurosci.* 43, 222-231.

Corradini, E., Babitt, J. L. and Lin, H. Y. (2009). The RGM/DRAGON Family of BMP Co-Receptors. *Cytokine Growth Factor Rev.* 20, 389-398.

Costa, M., Weir, M., Coulson, A., Sulston, J. and Kenyon, C. (1988). Posterior Pattern Formation in *C. elegans* Involves Position-Specific Expression of a Gene Containing a Homeobox. *Cell* 55, 747-756.

Cox, J. L., Mallanna, S. K., Luo, X. and Rizzino, A. (2010). Sox2 Uses Multiple Domains to Associate with Proteins Present in Sox2-Protein Complexes. *PLoS One* 5, e15486.

de Bont, J. M., Kros, J. M., Passier, M. M., Reddingius, R. E., Sillevis Smitt, P. A., Luider, T. M., den Boer, M. L. and Pieters, R. (2008). Differential Expression and Prognostic Significance of SOX Genes in Pediatric Medulloblastoma and Ependymoma Identified by Microarray Analysis. *Neuro Oncol.* 10, 648-660.

de Celis, J. F. and Barrio, R. (2009). Regulation and Function of Spalt Proteins during Animal Development. *Int. J. Dev. Biol.* 53, 1385-1398.

De Domenico, I., McVey Ward, D. and Kaplan, J. (2008a). Regulation of Iron Acquisition and Storage: Consequences for Iron-Linked Disorders. *Nat. Rev. Mol. Cell Biol.* 9, 72-81.

De Domenico, I., Nemeth, E., Nelson, J. M., Phillips, J. D., Ajioka, R. S., Kay, M. S., Kushner, J. P., Ganz, T., Ward, D. M. and Kaplan, J. (2008b). The Hepcidin-Binding Site on Ferroportin is Evolutionarily Conserved. *Cell. Metab.* 8, 146-156.

De Gobbi, M., Roetto, A., Piperno, A., Mariani, R., Alberti, F., Papanikolaou, G., Politou, M., Lockitch, G., Girelli, D., Fargion, S. et al. (2002). Natural History of Juvenile Haemochromatosis. *Br. J. Haematol.* 117, 973-979.

Deal, R. B. and Henikoff, S. (2011). The INTACT Method for Cell Type-Specific Gene Expression and Chromatin Profiling in Arabidopsis Thaliana. *Nat. Protoc.* 6, 56-68.

del Re, E., Babitt, J. L., Pirani, A., Schneyer, A. L. and Lin, H. Y. (2004). In the Absence of Type III Receptor, the Transforming Growth Factor (TGF)-Beta Type II-B Receptor Requires the Type I Receptor to Bind TGF-beta2. *J. Biol. Chem.* 279, 22765-22772.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988). A Genetic Pathway for the Development of the Caenorhabditis elegans HSN Motor Neurons. *Nature* 336, 638-646.

Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A. and Chinnaiyan, A. M. (2001). Delineation of Prognostic Biomarkers in Prostate Cancer. *Nature* 412, 822-826.

Du, X., She, E., Gelbart, T., Truksa, J., Lee, P., Xia, Y., Khovananth, K., Mudd, S., Mann, N., Moresco, E. M. et al. (2008). The Serine Protease TMPRSS6 is Required to Sense Iron Deficiency. *Science* 320, 1088-1092.

Duncan, E. L., Danoy, P., Kemp, J. P., Leo, P. J., McCloskey, E., Nicholson, G. C., Eastell, R., Prince, R. L., Eisman, J. A., Jones, G. et al. (2011). Genome-Wide Association Study using Extreme Truncate Selection Identifies Novel Genes Affecting Bone Mineral Density and Fracture Risk. *PLoS Genet.* 7, e1001372.

Dy, P., Penzo-Mendez, A., Wang, H., Pedraza, C. E., Macklin, W. B. and Lefebvre, V. (2008). The Three SoxC Proteins--Sox4, Sox11 and Sox12--Exhibit Overlapping Expression Patterns and Molecular Properties. *Nucleic Acids Res.* 36, 3101-3117.

Ernst, T., Hergenhahn, M., Kenzelmann, M., Cohen, C. D., Bonrouhi, M., Weninger, A., Klaren, R., Grone, E. F., Wiesel, M., Gudemann, C. et al. (2002). Decrease and Gain of Gene Expression are Equally Discriminatory Markers for Prostate Carcinoma: A Gene Expression Analysis on Total and Microdissected Prostate Tissue. *Am. J. Pathol.* 160, 2169-2180.

Estevez, M., Attisano, L., Wrana, J. L., Albert, P. S., Massague, J. and Riddle, D. L. (1993). The Daf-4 Gene Encodes a Bone Morphogenetic Protein Receptor Controlling *C. elegans* Dauer Larva Development. *Nature* 365, 644-649.

Finberg, K. E., Whittlesey, R. L., Fleming, M. D. and Andrews, N. C. (2010). Down-Regulation of Bmp/Smad Signaling by Tmprss6 is Required for Maintenance of Systemic Iron Homeostasis. *Blood* 115, 3817-3826.

Fitzgerald, D. P., Bradford, D. and Cooper, H. M. (2007). Neogenin is Expressed on Neurogenic and Gliogenic Progenitors in the Embryonic and Adult Central Nervous System. *Gene Expr. Patterns* 7, 784-792.

Fitzgerald, D. P., Cole, S. J., Hammond, A., Seaman, C. and Cooper, H. M. (2006a). Characterization of Neogenin-Expressing Neural Progenitor Populations and Migrating Neuroblasts in the Embryonic Mouse Forebrain. *Neuroscience* 142, 703-716.

Fitzgerald, D. P., Seaman, C. and Cooper, H. M. (2006b). Localization of Neogenin Protein during Morphogenesis in the Mouse Embryo. *Dev. Dyn.* 235, 1720-1725.

Foehr, M. L., Lindy, A. S., Fairbank, R. C., Amin, N. M., Xu, M., Yanowitz, J., Fire, A. Z. and Liu, J. (2006). An Antagonistic Role for the *C. elegans* Schnurri Homolog SMA-9 in Modulating TGFbeta Signaling during Mesodermal Patterning. *Development* 133, 2887-2896.

Foehr, M. L. and Liu, J. (2008). Dorsoventral Patterning of the *C. elegans* Postembryonic Mesoderm Requires both LIN-12/Notch and TGFbeta Signaling. *Dev. Biol.* 313, 256-266.

Folgueras, A. R., de Lara, F. M., Pendas, A. M., Garabaya, C., Rodriguez, F., Astudillo, A., Bernal, T., Cabanillas, R., Lopez-Otin, C. and Velasco, G. (2008). Membrane-Bound Serine Protease Matriptase-2 (Tmprss6) is an Essential Regulator of Iron Homeostasis. *Blood* 112, 2539-2545.

Friedman, R. S., Bangur, C. S., Zasloff, E. J., Fan, L., Wang, T., Watanabe, Y. and Kalos, M. (2004). Molecular and Immunological Evaluation of the Transcription Factor SOX-4 as a Lung Tumor Vaccine Antigen. *J. Immunol.* 172, 3319-3327.

Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M. and Jorgensen, E. M. (2008). Single-Copy Insertion of Transgenes in Caenorhabditis Elegans. *Nat. Genet.* 40, 1375-1383.

Fujikura, Y., Krijt, J. and Necas, E. (2011). Liver and Muscle Hemojuvelin are Differently Glycosylated. *BMC Biochem.* 12, 52.

Fukushige, T., Hawkins, M. G. and McGhee, J. D. (1998). The GATA-Factor Elt-2 is Essential for Formation of the Caenorhabditis Elegans Intestine. *Dev. Biol.* 198, 286-302.

Fung, W. Y., Fat, K. F., Eng, C. K. and Lau, C. K. (2007). Crm-1 Facilitates BMP Signaling to Control Body Size in Caenorhabditis Elegans. *Dev. Biol.* 311, 95-105.

Gad, J. M., Keeling, S. L., Wilks, A. F., Tan, S. S. and Cooper, H. M. (1997). The Expression Patterns of Guidance Receptors, DCC and Neogenin, are Spatially and Temporally Distinct Throughout Mouse Embryogenesis. *Dev. Biol.* 192, 258-273.

Gao, S., Steffen, J. and Laughon, A. (2005). Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. *J Biol Chem*. 280, 36158-64.

Geisbrecht, B. V., Dowd, K. A., Barfield, R. W., Longo, P. A. and Leahy, D. J. (2003). Netrin Binds Discrete Subdomains of DCC and UNC5 and Mediates Interactions between DCC and Heparin. *J. Biol. Chem.* 278, 32561-32568.

Georgi, L. L., Albert, P. S. and Riddle, D. L. (1990). Daf-1, a *C. elegans* Gene Controlling Dauer Larva Development, Encodes a Novel Receptor Protein Kinase. *Cell* 61, 635-645.

Gilleard, J. S., Shafi, Y., Barry, J. D. and McGhee, J. D. (1999). ELT-3: A Caenorhabditis Elegans GATA Factor Expressed in the Embryonic Epidermis during Morphogenesis. *Dev. Biol.* 208, 265-280.

Gomez, T. M. and Zheng, J. Q. (2006). The Molecular Basis for Calcium-Dependent Axon Pathfinding. *Nat. Rev. Neurosci.* 7, 115-125.

Gordon, K. J. and Blobe, G. C. (2008). Role of Transforming Growth Factor-Beta Superfamily Signaling Pathways in Human Disease. *Biochim. Biophys. Acta* 1782, 197-228.

Granato, M., Schnabel, H. and Schnabel, R. (1994). Pha-1, a Selectable Marker for Gene Transfer in *C. elegans*. *Nucleic Acids Res.* 22, 1762-1763.

Grant, K., Hanna-Rose, W. and Han, M. (2000). Sem-4 Promotes Vulval Cell-Fate Determination in Caenorhabditis Elegans through Regulation of Lin-39 Hox. *Dev. Biol.* 224, 496-506.

Greenwald, I. S., Sternberg, P. W. and Horvitz, H. R. (1983). The Lin-12 Locus Specifies Cell Fates in Caenorhabditis elegans. *Cell* 34, 435-444.

Grisaru, S., Cano-Gauci, D., Tee, J., Filmus, J. and Rosenblum, N. D. (2001). Glypican-3 Modulates BMP- and FGF-Mediated Effects during Renal Branching Morphogenesis. *Dev. Biol.* 231, 31-46.

Groppe, J., Greenwald, J., Wiater, E., Rodriguez-Leon, J., Economides, A. N., Kwiatkowski, W., Affolter, M., Vale, W. W., Belmonte, J. C. and Choe, S. (2002). Structural Basis of BMP Signalling Inhibition by the Cystine Knot Protein Noggin. *Nature* 420, 636-642.

Gumienny, T. L., Macneil, L., Zimmerman, C. M., Wang, H., Chin, L., Wrana, J. L. and Padgett, R. W. (2010). Caenorhabditis Elegans SMA-10/LRIG is a

Conserved Transmembrane Protein that Enhances Bone Morphogenetic Protein Signaling. *PLoS Genet.* 6, e1000963.

Gumienny, T. L., MacNeil, L. T., Wang, H., de Bono, M., Wrana, J. L. and Padgett, R. W. (2007). Glypican LON-2 is a Conserved Negative Regulator of BMP-Like Signaling in Caenorhabditis Elegans. *Curr. Biol.* 17, 159-164.

Hagihara, M., Endo, M., Hata, K., Higuchi, C., Takaoka, K., Yoshikawa, H. and Yamashita, T. (2011). Neogenin, a Receptor for Bone Morphogenetic Proteins. *J. Biol. Chem.* 286, 5157-5165.

Hata, K., Fujitani, M., Yasuda, Y., Doya, H., Saito, T., Yamagishi, S., Mueller, B. K. and Yamashita, T. (2006). RGMa Inhibition Promotes Axonal Growth and Recovery After Spinal Cord Injury. *J. Cell Biol.* 173, 47-58.

Halbrooks, P. J., Ding, R., Wozney, J. M. and Bain, G. (2007). Role of RGM Coreceptors in Bone Morphogenetic Protein Signaling. *J. Mol. Signal.* 2, 4.

Han, C., Yan, D., Belenkaya, T. Y. and Lin, X. (2005). Drosophila Glypicans Dally and Dally-Like Shape the Extracellular Wingless Morphogen Gradient in the Wing Disc. *Development* 132, 667-679.

Harfe, B. D. and Fire, A. (1998). Muscle and Nerve-Specific Regulation of a Novel NK-2 Class Homeodomain Factor in Caenorhabditis elegans. *Development* 125, 421-429.

Harfe, B. D., Branda, C. S., Krause, M., Stern, M. J. and Fire, A. (1998). MyoD and the Specification of Muscle and Non-Muscle Fates during Postembryonic Development of the *C. elegans* Mesoderm. *Development* 125, 2479-2488.

Harfe, B. D., Vaz Gomes, A., Kenyon, C., Liu, J., Krause, M. and Fire, A. (1998). Analysis of a Caenorhabditis elegans Twist Homolog Identifies Conserved and Divergent Aspects of Mesodermal Patterning. *Genes Dev.* 12, 2623-2635.

Hargrave, M., Wright, E., Kun, J., Emery, J., Cooper, L. and Koopman, P. (1997). Expression of the Sox11 Gene in Mouse Embryos Suggests Roles in Neuronal Maturation and Epithelio-Mesenchymal Induction. *Dev. Dyn.* 210, 79-86.

Hata, K., Fujitani, M., Yasuda, Y., Doya, H., Saito, T., Yamagishi, S., Mueller, B. K. and Yamashita, T. (2006). RGMa Inhibition Promotes Axonal Growth and Recovery After Spinal Cord Injury. *J. Cell Biol.* 173, 47-58.

Hide, T., Takezaki, T., Nakatani, Y., Nakamura, H., Kuratsu, J. and Kondo, T. (2009). Sox11 Prevents Tumorigenesis of Glioma-Initiating Cells by Inducing Neuronal Differentiation. *Cancer Res.* 69, 7953-7959.

Hong, K., Hinck, L., Nishiyama, M., Poo, M. M., Tessier-Lavigne, M. and Stein, E. (1999). A Ligand-Gated Association between Cytoplasmic Domains of UNC5 and DCC Family Receptors Converts Netrin-Induced Growth Cone Attraction to Repulsion. *Cell* 97, 927-941.

Hong, M., Schachter, K. A., Jiang, G. and Krauss, R. S. (2012). Neogenin Regulates Sonic Hedgehog Pathway Activity during Digit Patterning. *Dev. Dyn.* 241, 627-637.

Hoser, M., Baader, S. L., Bosl, M. R., Ihmer, A., Wegner, M. and Sock, E. (2007). Prolonged Glial Expression of Sox4 in the CNS Leads to Architectural Cerebellar Defects and Ataxia. *J. Neurosci.* 27, 5495-5505.

Hoser, M., Potzner, M. R., Koch, J. M., Bosl, M. R., Wegner, M. and Sock, E. (2008). Sox12 Deletion in the Mouse Reveals Nonreciprocal Redundancy with the Related Sox4 and Sox11 Transcription Factors. *Mol. Cell. Biol.* 28, 4675-4687.

Huang, F. W., Pinkus, J. L., Pinkus, G. S., Fleming, M. D. and Andrews, N. C. (2005). A Mouse Model of Juvenile Hemochromatosis. *J. Clin. Invest.* 115, 2187-2191.

Itokazu, T., Fujita, Y., Takahashi, R. and Yamashita, T. (2012). Identification of the Neogenin-Binding Site on the Repulsive Guidance Molecule a. *PLoS One* 7, e32791.

Jarjour, A. A., Bull, S. J., Almasieh, M., Rajasekharan, S., Baker, K. A., Mui, J., Antel, J. P., Di Polo, A. and Kennedy, T. E. (2008). Maintenance of Axo-Oligodendroglial Paranodal Junctions Requires DCC and Netrin-1. *J. Neurosci.* 28, 11003-11014.

Jauch, R., Ng, C. K., Narasimhan, K. and Kolatkar, P. R. (2012). The Crystal Structure of the Sox4 HMG Domain-DNA Complex Suggests a Mechanism for Positional Interdependence in DNA Recognition. *Biochem. J.* 443, 39-47.

Jiang, Y., Horner, V. and Liu, J. (2005). The HMX Homeodomain Protein MLS-2 Regulates Cleavage Orientation, Cell Proliferation and Cell Fate Specification in the *C. elegans* Postembryonic Mesoderm. *Development* 132, 4119-4130.

Jiang, Y., Shi, H., Amin, N. M., Sultan, I. and Liu, J. (2008). Mesodermal Expression of the *C. elegans* HMX Homolog Mls-2 Requires the PBC Homolog CEH-20. *Mech. Dev.* 125, 451-461.

Jiang, Y., Shi, H. and Liu, J. (2009). Two Hox Cofactors, the Meis/Hth Homolog UNC-62 and the Pbx/Exd Homolog CEH-20, Function Together during *C. elegans* Postembryonic Mesodermal Development. *Dev. Biol.* 334, 535-546.

Johnson, K., Kirkpatrick, H., Comer, A., Hoffmann, F. M. and Laughon, A. (1999). Interaction of Smad Complexes with Tripartite DNA-Binding Sites. *J. Biol. Chem.* 274, 20709-20716.

Kang, J. S., Yi, M. J., Zhang, W., Feinleib, J. L., Cole, F. and Krauss, R. S. (2004). Netrins and Neogenin Promote Myotube Formation. *J. Cell Biol.* 167, 493-504.

Kamath, R. S. and Ahringer, J. (2003). Genome-Wide RNAi Screening in Caenorhabditis elegans. *Methods* 30, 313-321.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M. et al. (2003). Systematic Functional Analysis of the Caenorhabditis elegans Genome using RNAi. *Nature* 421, 231-237.

Kanomata, K., Kokabu, S., Nojima, J., Fukuda, T. and Katagiri, T. (2009). DRAGON, a GPI-Anchored Membrane Protein, Inhibits BMP Signaling in C2C12 Myoblasts. *Genes Cells* 14, 695-702.

Karantzali, E., Lekakis, V., Ioannou, M., Hadjimichael, C., Papamatheakis, J. and Kretsovali, A. (2011). Sall1 Regulates Embryonic Stem Cell Differentiation in Association with Nanog. *J. Biol. Chem.* 286, 1037-1045.

Kautz, L., Meynard, D., Monnier, A., Darnaud, V., Bouvet, R., Wang, R. H., Deng, C., Vaulont, S., Mosser, J., Coppin, H. et al. (2008a). Iron Regulates Phosphorylation of Smad1/5/8 and Gene Expression of Bmp6, Smad7, Id1, and Atoh8 in the Mouse Liver. *Blood* 112, 1503-1509.

Kautz, L., Meynard, D., Monnier, A., Darnaud, V., Bouvet, R., Wang, R. H., Deng, C., Vaulont, S., Mosser, J., Coppin, H. et al. (2008b). Iron Regulates Phosphorylation of Smad1/5/8 and Gene Expression of Bmp6, Smad7, Id1, and Atoh8 in the Mouse Liver. *Blood* 112, 1503-1509.

Kee, N., Wilson, N., De Vries, M., Bradford, D., Key, B. and Cooper, H. M. (2008). Neogenin and RGMa Control Neural Tube Closure and Neuroepithelial Morphology by Regulating Cell Polarity. *J. Neurosci.* 28, 12643-12653.

Keeling, S. L., Gad, J. M. and Cooper, H. M. (1997). Mouse Neogenin, a DCC-Like Molecule, has Four Splice Variants and is Expressed Widely in the Adult Mouse and during Embryogenesis. *Oncogene* 15, 691-700.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E. D., Chan, S. S., Culotti, J. G. and Tessier-Lavigne, M. (1996). Deleted in Colorectal Cancer (DCC) Encodes a Netrin Receptor. *Cell* 87, 175-185.

Kennedy, T. E., Serafini, T., de la Torre, J. R. and Tessier-Lavigne, M. (1994). Netrins are Diffusible Chemotropic Factors for Commissural Axons in the Embryonic Spinal Cord. *Cell* 78, 425-435.

Kidd, A. R.,3rd, Miskowski, J. A., Siegfried, K. R., Sawa, H. and Kimble, J. (2005). A Beta-Catenin Identified by Functional rather than Sequence Criteria and its Role in Wnt/MAPK Signaling. *Cell* 121, 761-772.

Kiefer, S. M., McDill, B. W., Yang, J. and Rauchman, M. (2002). Murine Sall1 Represses Transcription by Recruiting a Histone Deacetylase Complex. *J. Biol. Chem.* 277, 14869-14876.

Kirilly, D., Gu, Y., Huang, Y., Wu, Z., Bashirullah, A., Low, B. C., Kolodkin, A. L., Wang, H. and Yu, F. (2009). A Genetic Pathway Composed of Sox14 and Mical Governs Severing of Dendrites during Pruning. *Nat. Neurosci.* 12, 1497-1505.

Knoepfler, P. S., Lu, Q. and Kamps, M. P. (1996). Pbx-1 Hox Heterodimers Bind DNA on Inseparable Half-Sites that Permit Intrinsic DNA Binding Specificity of

the Hox Partner at Nucleotides 3' to a TAAT Motif. *Nucleic Acids Res.* 24, 2288-2294.

Koeberle, P. D., Tura, A., Tassew, N. G., Schlichter, L. C. and Monnier, P. P. (2010). The Repulsive Guidance Molecule, RGMa, Promotes Retinal Ganglion Cell Survival *in vitro* and *in vivo*. *Neuroscience* 169, 495-504.

Kostas, S. A. and Fire, A. (2002). The T-Box Factor MLS-1 Acts as a Molecular Switch during Specification of Nonstriated Muscle in *C. elegans*. *Genes Dev.* 16, 257-269.

Koh, K., Peyrot, S. M., Wood, C. G., Wagmaister, J. A., Maduro, M. F., Eisenmann, D. M. and Rothman, J. H. (2002). Cell Fates and Fusion in the *C. elegans* Vulval Primordium are Regulated by the EGL-18 and ELT-6 GATA Factors -- Apparent Direct Targets of the LIN-39 Hox Protein. *Development* 129, 5171-5180.

Kramer, J. M. and Johnson, J. J. (1993). Analysis of Mutations in the Sqt-1 and Rol-6 Collagen Genes of Caenorhabditis Elegans. *Genetics* 135, 1035-1045.

Kretzschmar, M., Doody, J. and Massague, J. (1997). Opposing BMP and EGF Signalling Pathways Converge on the TGF-Beta Family Mediator Smad1. *Nature* 389, 618-622.

Kretzschmar, M., Doody, J., Timokhina, I. and Massague, J. (1999). A Mechanism of Repression of TGFbeta/ Smad Signaling by Oncogenic Ras. *Genes Dev.* 13, 804-816.

Krijt, J., Fujikura, Y., Ramsay, A. J., Velasco, G. and Necas, E. (2011). Liver Hemojuvelin Protein Levels in Mice Deficient in Matriptase-2 (Tmprss6). *Blood Cells Mol. Dis.* 47, 133-137.

Krishna, S., Maduzia, L. L. and Padgett, R. W. (1999). Specificity of TGFbeta Signaling is Conferred by Distinct Type I Receptors and their Associated SMAD Proteins in Caenorhabditis Elegans. *Development* 126, 251-260.

Kruger, R. P., Lee, J., Li, W. and Guan, K. L. (2004). Mapping Netrin Receptor Binding Reveals Domains of Unc5 Regulating its Tyrosine Phosphorylation. *J. Neurosci.* 24, 10826-10834.

Kubo, T., Endo, M., Hata, K., Taniguchi, J., Kitajo, K., Tomura, S., Yamaguchi, A., Mueller, B. K. and Yamashita, T. (2008). Myosin IIA is Required for Neurite Outgrowth Inhibition Produced by Repulsive Guidance Molecule. *J. Neurochem.* 105, 113-126.

Kuhlbrodt, K., Herbarth, B., Sock, E., Enderich, J., Hermans-Borgmeyer, I. and Wegner, M. (1998). Cooperative Function of POU Proteins and SOX Proteins in Glial Cells. *J. Biol. Chem.* 273, 16050-16057.

Kuninger, D., Kuns-Hashimoto, R., Kuzmickas, R. and Rotwein, P. (2006a). Complex Biosynthesis of the Muscle-Enriched Iron Regulator RGMc. *J. Cell. Sci.* 119, 3273-3283.

Kuninger, D., Kuns-Hashimoto, R., Kuzmickas, R. and Rotwein, P. (2006b). Complex Biosynthesis of the Muscle-Enriched Iron Regulator RGMc. *J. Cell. Sci.* 119, 3273-3283.

Kuninger, D., Kuzmickas, R., Peng, B., Pintar, J. E. and Rotwein, P. (2004). Gene Discovery by Microarray: Identification of Novel Genes Induced during Growth Factor-Mediated Muscle Cell Survival and Differentiation. *Genomics* 84, 876-889.

Kuns-Hashimoto, R., Kuninger, D., Nili, M. and Rotwein, P. (2008). Selective Binding of RGMc/hemojuvelin, a Key Protein in Systemic Iron Metabolism, to BMP-2 and Neogenin. *Am. J. Physiol. Cell. Physiol.* 294, C994-C1003.

Lah, G. J. and Key, B. (2012). Novel Roles of the Chemorepellent Axon Guidance Molecule RGMa in Cell Migration and Adhesion. *Mol. Cell. Biol.* 32, 968-980.

Lai Wing Sun, K., Correia, J. P. and Kennedy, T. E. (2011). Netrins: Versatile Extracellular Cues with Diverse Functions. *Development* 138, 2153-2169.

Lai, Y. H., Cheng, J., Cheng, D., Feasel, M. E., Beste, K. D., Peng, J., Nusrat, A. and Moreno, C. S. (2011). SOX4 Interacts with Plakoglobin in a Wnt3a-Dependent Manner in Prostate Cancer Cells. *BMC Cell Biol.* 12, 50.

Lapointe, J., Li, C., Higgins, J. P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U. et al. (2004). Gene Expression Profiling Identifies Clinically Relevant Subtypes of Prostate Cancer. *Proc. Natl. Acad. Sci. U. S. A.* 101, 811-816.

Lee, C. J., Appleby, V. J., Orme, A. T., Chan, W. I. and Scotting, P. J. (2002). Differential Expression of SOX4 and SOX11 in Medulloblastoma. *J. Neurooncol.* 57, 201-214.

Lee, D. H., Zhou, L. J., Zhou, Z., Xie, J. X., Jung, J. U., Liu, Y., Xi, C. X., Mei, L. and Xiong, W. C. (2010). Neogenin Inhibits HJV Secretion and Regulates BMP-Induced Hepcidin Expression and Iron Homeostasis. *Blood* 115, 3136-3145.

Lejmi, E., Leconte, L., Pedron-Mazoyer, S., Ropert, S., Raoul, W., Lavalette, S., Bouras, I., Feron, J. G., Maitre-Boube, M., Assayag, F. et al. (2008). Netrin-4 Inhibits Angiogenesis Via Binding to Neogenin and Recruitment of Unc5B. *Proc. Natl. Acad. Sci. U. S. A.* 105, 12491-12496.

Lewis, K. A., Gray, P. C., Blount, A. L., MacConell, L. A., Wiater, E., Bilezikjian, L. M. and Vale, W. (2000). Betaglycan Binds Inhibin and can Mediate Functional Antagonism of Activin Signalling. *Nature* 404, 411-414.

Liang, J., Lints, R., Foehr, M. L., Tokarz, R., Yu, L., Emmons, S. W., Liu, J. and Savage-Dunn, C. (2003). The Caenorhabditis Elegans Schnurri Homolog Sma-9 Mediates Stage- and Cell Type-Specific Responses to DBL-1 BMP-Related Signaling. *Development* 130, 6453-6464.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S. Y., Qin, J., Wong, J., Cooney, A. J., Liu, D. et al. (2008). Nanog and Oct4 Associate with Unique Transcriptional Repression Complexes in Embryonic Stem Cells. *Nat. Cell Biol.* 10, 731-739.

Li, J., Ye, L., Kynaston, H. G. and Jiang, W. G. (2012). Repulsive Guidance Molecules, Novel Bone Morphogenetic Protein Co-Receptors, are Key Regulators of the Growth and Aggressiveness of Prostate Cancer Cells. *Int. J. Oncol.* 40, 544-550.

Li, J., Ye, L., Mansel, R. E. and Jiang, W. G. (2011). Potential Prognostic Value of Repulsive Guidance Molecules in Breast Cancer. *Anticancer Res.* 31, 1703-1711.

Li, J., Ye, L., Sanders, A. J. and Jiang, W. G. (2012). Repulsive Guidance Molecule B (RGMB) Plays Negative Roles in Breast Cancer by Coordinating BMP Signaling. *J. Cell. Biochem.*

Li, W., Lee, J., Vikis, H. G., Lee, S. H., Liu, G., Aurandt, J., Shen, T. L., Fearon, E. R., Guan, J. L., Han, M. et al. (2004). Activation of FAK and Src are Receptor-Proximal Events Required for Netrin Signaling. *Nat. Neurosci.* 7, 1213-1221.

Lin, A. C. and Holt, C. E. (2007). Local Translation and Directional Steering in Axons. *EMBO J.* 26, 3729-3736.

Lin, L., Goldberg, Y. P. and Ganz, T. (2005a). Competitive Regulation of Hepcidin mRNA by Soluble and Cell-Associated Hemojuvelin. *Blood* 106, 2884-2889.

Lin, L., Goldberg, Y. P. and Ganz, T. (2005b). Competitive Regulation of Hepcidin mRNA by Soluble and Cell-Associated Hemojuvelin. *Blood* 106, 2884-2889.

Lin, L., Lee, V. M., Wang, Y., Lin, J. S., Sock, E., Wegner, M. and Lei, L. (2011). Sox11 Regulates Survival and Axonal Growth of Embryonic Sensory Neurons. *Dev. Dyn.* 240, 52-64.

Lin, L., Nemeth, E., Goodnough, J. B., Thapa, D. R., Gabayan, V. and Ganz, T. (2008). Soluble Hemojuvelin is Released by Proprotein Convertase-Mediated Cleavage at a Conserved Polybasic RNRR Site. *Blood Cells Mol. Dis.* 40, 122-131.

Lin, L., Valore, E. V., Nemeth, E., Goodnough, J. B., Gabayan, V. and Ganz, T. (2007). Iron Transferrin Regulates Hepcidin Synthesis in Primary Hepatocyte Culture through Hemojuvelin and BMP2/4. *Blood* 110, 2182-2189.

Liu, G., Beggs, H., Jurgensen, C., Park, H. T., Tang, H., Gorski, J., Jones, K. R., Reichardt, L. F., Wu, J. and Rao, Y. (2004). Netrin Requires Focal Adhesion

Kinase and Src Family Kinases for Axon Outgrowth and Attraction. *Nat. Neurosci.* 7, 1222-1232.

Liu, H., Strauss, T. J., Potts, M. B. and Cameron, S. (2006). Direct Regulation of Egl-1 and of Programmed Cell Death by the Hox Protein MAB-5 and by CEH-20, a *C. elegans* Homolog of Pbx1. *Development* 133, 641-650.

Liu, J. and Fire, A. (2000). Overlapping Roles of Two Hox Genes and the Exd Ortholog Ceh-20 in Diversification of the *C. elegans* Postembryonic Mesoderm. *Development* 127, 5179-5190.

Liu, J., Phillips, B. T., Amaya, M. F., Kimble, J. and Xu, W. (2008). The *C. elegans* SYS-1 Protein is a Bona Fide Beta-Catenin. *Dev. Cell.* 14, 751-761.

Liu, P., Ramachandran, S., Ali Seyed, M., Scharer, C. D., Laycock, N., Dalton, W. B., Williams, H., Karanam, S., Datta, M. W., Jaye, D. L. et al. (2006). Sex-Determining Region Y Box 4 is a Transforming Oncogene in Human Prostate Cancer Cells. *Cancer Res.* 66, 4011-4019.

Liu, X., Hashimoto, M., Horii, H., Yamaguchi, A., Naito, K. and Yamashita, T. (2009). Repulsive Guidance Molecule b Inhibits Neurite Growth and is Increased After Spinal Cord Injury. *Biochem. Biophys. Res. Commun.* 382, 795-800.

Liu, Y., Peng, Y., Dai, P. G., Du, Q. S., Mei, L. and Xiong, W. C. (2012). Differential Regulation of Myosin X Movements by its Cargos, DCC and Neogenin. *J. Cell. Sci.* 125, 751-762.

Ma, C. H., Brenner, G. J., Omura, T., Samad, O. A., Costigan, M., Inquimbert, P., Niederkofler, V., Salie, R., Sun, C. C., Lin, H. Y. et al. (2011). The BMP Coreceptor RGMb Promotes while the Endogenous BMP Antagonist Noggin Reduces Neurite Outgrowth and Peripheral Nerve Regeneration by Modulating BMP Signaling. *J. Neurosci.* 31, 18391-18400.

Maduzia, L. L., Gumienny, T. L., Zimmerman, C. M., Wang, H., Shetgiri, P., Krishna, S., Roberts, A. F. and Padgett, R. W. (2002). Lon-1 Regulates Caenorhabditis Elegans Body Size Downstream of the Dbl-1 TGF Beta Signaling Pathway. *Dev. Biol.* 246, 418-428.

Maeda, I., Kohara, Y., Yamamoto, M. and Sugimoto, A. (2001). Large-Scale Analysis of Gene Function in Caenorhabditis elegans by High-Throughput RNAi. *Curr. Biol.* 11, 171-176.

Magee, J. A., Araki, T., Patil, S., Ehrig, T., True, L., Humphrey, P. A., Catalona, W. J., Watson, M. A. and Milbrandt, J. (2001). Expression Profiling Reveals Hepsin Overexpression in Prostate Cancer. *Cancer Res.* 61, 5692-5696.

Mann, R. S. and Affolter, M. (1998). Hox Proteins Meet More Partners. *Curr. Opin. Genet. Dev.* 8, 423-429.

Massague, J. (1992). Receptors for the TGF-Beta Family. *Cell* 69, 1067-1070.

Massague, J. (1998). TGF-Beta Signal Transduction. *Annu. Rev. Biochem.* 67, 753-791.

Massague, J. (2008). TGFbeta in Cancer. *Cell* 134, 215-230.

Massague, J., Blain, S. W. and Lo, R. S. (2000a). TGFbeta Signaling in Growth Control, Cancer, and Heritable Disorders. *Cell* 103, 295-309.

Massague, J., Blain, S. W. and Lo, R. S. (2000b). TGFbeta Signaling in Growth Control, Cancer, and Heritable Disorders. *Cell* 103, 295-309.

Massague, J. and Chen, Y. G. (2000). Controlling TGF-Beta Signaling. *Genes Dev.* 14, 627-644.

Massague, J., Seoane, J. and Wotton, D. (2005a). Smad Transcription Factors. *Genes Dev.* 19, 2783-2810.

Massague, J., Seoane, J. and Wotton, D. (2005b). Smad Transcription Factors. *Genes Dev.* 19, 2783-2810.

Massague, J. and Wotton, D. (2000a). Transcriptional Control by the TGF-beta/Smad Signaling System. *EMBO J.* 19, 1745-1754.

Massague, J. and Wotton, D. (2000b). Transcriptional Control by the TGF-beta/Smad Signaling System. *EMBO J.* 19, 1745-1754.

Matsunaga, E., Nakamura, H. and Chedotal, A. (2006). Repulsive Guidance Molecule Plays Multiple Roles in Neuronal Differentiation and Axon Guidance. *J. Neurosci.* 26, 6082-6088.

Matsunaga, E., Tauszig-Delamasure, S., Monnier, P. P., Mueller, B. K., Strittmatter, S. M., Mehlen, P. and Chedotal, A. (2004). RGM and its Receptor Neogenin Regulate Neuronal Survival. *Nat. Cell Biol.* 6, 749-755.

Matsuura, I., Chiang, K. N., Lai, C. Y., He, D., Wang, G., Ramkumar, R., Uchida, T., Ryo, A., Lu, K. and Liu, F. (2010). Pin1 Promotes Transforming Growth Factor-Beta-Induced Migration and Invasion. *J. Biol. Chem.* 285, 1754-1764.

Mavropoulos, A., Devos, N., Biemar, F., Zecchin, E., Argenton, F., Edlund, H., Motte, P., Martial, J. A. and Peers, B. (2005). Sox4b is a Key Player of Pancreatic Alpha Cell Differentiation in Zebrafish. *Dev. Biol.* 285, 211-223.

Maxson, J. E., Enns, C. A. and Zhang, A. S. (2009). Processing of Hemojuvelin Requires Retrograde Trafficking to the Golgi in HepG2 Cells. *Blood* 113, 1786-1793.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient Gene Transfer in *C. elegans*: Extrachromosomal Maintenance and Integration of Transforming Sequences. *EMBO J.* 10, 3959-3970.

Medina, P. P., Castillo, S. D., Blanco, S., Sanz-Garcia, M., Largo, C., Alvarez, S., Yokota, J., Gonzalez-Neira, A., Benitez, J., Clevers, H. C. et al. (2009). The SRY-HMG Box Gene, SOX4, is a Target of Gene Amplification at Chromosome 6p in Lung Cancer. *Hum. Mol. Genet.* 18, 1343-1352.

Meynard, D., Kautz, L., Darnaud, V., Canonne-Hergaux, F., Coppin, H. and Roth, M. P. (2009). Lack of the Bone Morphogenetic Protein BMP6 Induces Massive Iron Overload. *Nat. Genet.* 41, 478-481.

Meynard, D., Vaja, V., Sun, C. C., Corradini, E., Chen, S., Lopez-Otin, C., Grgurevic, L., Hong, C. C., Stirnberg, M., Gutschow, M. et al. (2011). Regulation of TMPRSS6 by BMP6 and Iron in Human Cells and Mice. *Blood* 118, 747-756.

Mirakaj, V., Brown, S., Laucher, S., Steinl, C., Klein, G., Kohler, D., Skutella, T., Meisel, C., Brommer, B., Rosenberger, P. et al. (2011). Repulsive Guidance Molecule-A (RGM-A) Inhibits Leukocyte Migration and Mitigates Inflammation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 6555-6560.

Miyazono K, Hellman U, Wernstedt C, Heldin CH. (1988). Latent high molecular weight complex of transforming growth factor beta 1. Purification from human platelets and structural characterization. *J Biol Chem.* 263:6407–6415.

Moreno, C. S. (2010). The Sex-Determining Region Y-Box 4 and Homeobox C6 Transcriptional Networks in Prostate Cancer Progression: Crosstalk with the Wnt, Notch, and PI3K Pathways. *Am. J. Pathol.* 176, 518-527.

Moore, S. W., Tessier-Lavigne, M. and Kennedy, T. E. (2007). Netrins and their Receptors. *Adv. Exp. Med. Biol.* 621, 17-31.

Monnier, P. P., Sierra, A., Macchi, P., Deitinghoff, L., Andersen, J. S., Mann, M., Flad, M., Hornberger, M. R., Stahl, B., Bonhoeffer, F. et al. (2002). RGM is a Repulsive Guidance Molecule for Retinal Axons. *Nature* 419, 392-395.

Morita, K., Chow, K. L. and Ueno, N. (1999). Regulation of Body Length and Male Tail Ray Pattern Formation of Caenorhabditis Elegans by a Member of TGF-Beta Family. *Development* 126, 1337-1347.

Morita, K., Flemming, A. J., Sugihara, Y., Mochii, M., Suzuki, Y., Yoshida, S., Wood, W. B., Kohara, Y., Leroi, A. M. and Ueno, N. (2002). A Caenorhabditis Elegans TGF-Beta, DBL-1, Controls the Expression of LON-1, a PR-Related Protein, that Regulates Polyploidization and Body Length. *EMBO J.* 21, 1063-1073.

Moustakas, A. and Heldin, C. H. (2009). The Regulation of TGFbeta Signal Transduction. *Development* 136, 3699-3714.

Mu, L., Berti, L., Masserdotti, G., Covic, M., Michaelidis, T. M., Doberauer, K., Merz, K., Rehfeld, F., Haslinger, A., Wegner, M. et al. (2012). SoxC Transcription Factors are Required for Neuronal Differentiation in Adult Hippocampal Neurogenesis. *J. Neurosci.* **32,** 3067-3080.

Muramatsu, R., Kubo, T., Mori, M., Nakamura, Y., Fujita, Y., Akutsu, T., Okuno, T., Taniguchi, J., Kumanogoh, A., Yoshida, M. et al. (2011). RGMa Modulates T Cell Responses and is Involved in Autoimmune Encephalomyelitis. *Nat. Med.* 17, 488-494.

Nagayoshi, Y., Nakayama, M., Suzuki, S., Hokamaki, J., Shimomura, H., Tsujita, K., Fukuda, M., Yamashita, T., Nakamura, Y., Sugiyama, S. et al. (2008). A Q312X Mutation in the Hemojuvelin Gene is Associated with Cardiomyopathy due to Juvenile Haemochromatosis. *Eur. J. Heart Fail.* 10, 1001-1006.

Nash, B., Colavita, A., Zheng, H., Roy, P. J. and Culotti, J. G. (2000). The Forkhead Transcription Factor UNC-130 is Required for the Graded Spatial Expression of the UNC-129 TGF-Beta Guidance Factor in *C. elegans*. *Genes Dev.* 14, 2486-2500.

Neff, A. W., King, M. W. and Mescher, A. L. (2011). Dedifferentiation and the Role of sall4 in Reprogramming and Patterning during Amphibian Limb Regeneration. *Dev. Dyn.* 240, 979-989.

Niederkofler, V., Salie, R., Sigrist, M. and Arber, S. (2004). Repulsive Guidance Molecule (RGM) Gene Function is Required for Neural Tube Closure but Not Retinal Topography in the Mouse Visual System. *J. Neurosci.* 24, 808-818.

Nili, M., Shinde, U. and Rotwein, P. (2010). Soluble Repulsive Guidance Molecule c/hemojuvelin is a Broad Spectrum Bone Morphogenetic Protein (BMP) Antagonist and Inhibits both BMP2- and BMP6-Mediated Signaling and Gene Expression. *J. Biol. Chem.* 285, 24783-24792.

Nissen-Meyer, L. S., Jemtland, R., Gautvik, V. T., Pedersen, M. E., Paro, R., Fortunati, D., Pierroz, D. D., Stadelmann, V. A., Reppe, S., Reinholt, F. P. et al.

(2007). Osteopenia, Decreased Bone Formation and Impaired Osteoblast Development in Sox4 Heterozygous Mice. *J. Cell. Sci.* 120, 2785-2795.

Nohra, R., Beyeen, A. D., Guo, J. P., Khademi, M., Sundqvist, E., Hedreul, M. T., Sellebjerg, F., Smestad, C., Oturai, A. B., Harbo, H. F. et al. (2010). RGMA and IL21R show Association with Experimental Inflammation and Multiple Sclerosis. *Genes Immun.* 11, 279-293.

O'Connor, M. B., Umulis, D., Othmer, H. G. and Blair, S. S. (2006). Shaping BMP Morphogen Gradients in the Drosophila Embryo and Pupal Wing. *Development* 133, 183-193.

Oikawa, T., Kamiya, A., Kakinuma, S., Zeniya, M., Nishinakamura, R., Tajiri, H. and Nakauchi, H. (2009). Sall4 Regulates Cell Fate Decision in Fetal Hepatic stem/progenitor Cells. *Gastroenterology* 136, 1000-1011.

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A. and Fire, A. (1993). Sequence Requirements for Myosin Gene Expression and Regulation in Caenorhabditis Elegans. *Genetics* 135, 385-404.

Pagani, A., Silvestri, L., Nai, A. and Camaschella, C. (2008). Hemojuvelin N-Terminal Mutants Reach the Plasma Membrane but do Not Activate the Hepcidin Response. *Haematologica* 93, 1466-1472.

Pan, X., Zhao, J., Zhang, W. N., Li, H. Y., Mu, R., Zhou, T., Zhang, H. Y., Gong, W. L., Yu, M., Man, J. H. et al. (2009a). Induction of SOX4 by DNA Damage is

Critical for p53 Stabilization and Function. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3788-3793.

Pan, X., Zhao, J., Zhang, W. N., Li, H. Y., Mu, R., Zhou, T., Zhang, H. Y., Gong, W. L., Yu, M., Man, J. H. et al. (2009b). Induction of SOX4 by DNA Damage is Critical for p53 Stabilization and Function. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3788-3793.

Papanikolaou, G., Samuels, M. E., Ludwig, E. H., MacDonald, M. L., Franchini, P. L., Dube, M. P., Andres, L., MacFarlane, J., Sakellaropoulos, N., Politou, M. et al. (2004). Mutations in HFE2 Cause Iron Overload in Chromosome 1q-Linked Juvenile Hemochromatosis. *Nat. Genet.* 36, 77-82.

Park, C. H., Valore, E. V., Waring, A. J. and Ganz, T. (2001). Hepcidin, a Urinary Antimicrobial Peptide Synthesized in the Liver. *J. Biol. Chem.* 276, 7806-7810.

Park, K. W., Crouse, D., Lee, M., Karnik, S. K., Sorensen, L. K., Murphy, K. J., Kuo, C. J. and Li, D. Y. (2004). The Axonal Attractant Netrin-1 is an Angiogenic Factor. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16210-16215.

Partridge, F. A., Gravato-Nobre, M. J. and Hodgkin, J. (2010). Signal Transduction Pathways that Function in both Development and Innate Immunity. *Dev. Dyn.* 2010 Feb 3. [Epub ahead of print].

Patterson, G. I. and Padgett, R. W. (2000). TGF Beta-Related Pathways. Roles in Caenorhabditis Elegans Development. *Trends Genet.* 16, 27-33.

Penzo-Mendez, A. I. (2010). Critical Roles for SoxC Transcription Factors in Development and Cancer. *Int. J. Biochem. Cell Biol.* 42, 425-428.

Phochanukul, N. and Russell, S. (2010). No Backbone but Lots of Sox: Invertebrate Sox Genes. *Int. J. Biochem. Cell Biol.* 42, 453-464.

Pigeon, C., Ilyin, G., Courselaud, B., Leroyer, P., Turlin, B., Brissot, P. and Loreal, O. (2001). A New Mouse Liver-Specific Gene, Encoding a Protein Homologous to Human Antimicrobial Peptide Hepcidin, is Overexpressed during Iron Overload. *J. Biol. Chem.* 276, 7811-7819.

Potts, M. B., Wang, D. P. and Cameron, S. (2009). Trithorax, Hox, and TALE-Class Homeodomain Proteins Ensure Cell Survival through Repression of the BH3-Only Gene Egl-1. *Dev. Biol.* 329, 374-385.

Potzner, M. R., Griffel, C., Lutjen-Drecoll, E., Bosl, M. R., Wegner, M. and Sock, E. (2007). Prolonged Sox4 Expression in Oligodendrocytes Interferes with Normal Myelination in the Central Nervous System. *Mol. Cell. Biol.* 27, 5316-5326.

Potzner, M. R., Tsarovina, K., Binder, E., Penzo-Mendez, A., Lefebvre, V., Rohrer, H., Wegner, M. and Sock, E. (2010). Sequential Requirement of Sox4 and Sox11 during Development of the Sympathetic Nervous System. *Development* 137, 775-784.

Pramoonjago, P., Baras, A. S. and Moskaluk, C. A. (2006). Knockdown of Sox4 Expression by RNAi Induces Apoptosis in ACC3 Cells. *Oncogene* 25, 5626-5639.

Pyrowolakis, G., Hartmann, B., Müller, B., Basler, K. and Affolter, M. (2004). A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. *Dev Cell.* 7, 229-40.

Qadota, H., Inoue, M., Hikita, T., Koppen, M., Hardin, J. D., Amano, M., Moerman, D. G. and Kaibuchi, K. (2007). Establishment of a Tissue-Specific RNAi System in *C. elegans*. *Gene* 400, 166-173.

Qin, S., Yu, L., Gao, Y., Zhou, R. and Zhang, C. (2007). Characterization of the Receptors for Axon Guidance Factor Netrin-4 and Identification of the Binding Domains. *Mol. Cell. Neurosci.* 34, 243-250.

Rajasekharan, S. and Kennedy, T. E. (2009). The Netrin Protein Family. *Genome Biol.* 10, 239.

Ramsay, A. J., Hooper, J. D., Folgueras, A. R., Velasco, G. and Lopez-Otin, C. (2009). Matriptase-2 (TMPRSS6): A Proteolytic Regulator of Iron Homeostasis. *Haematologica* 94, 840-849.

Rao, S., Zhen, S., Roumiantsev, S., McDonald, L. T., Yuan, G. C. and Orkin, S. H. (2010). Differential Roles of Sall4 Isoforms in Embryonic Stem Cell Pluripotency. *Mol. Cell. Biol.* 30, 5364-5380.

Ren, P., Lim, C. S., Johnsen, R., Albert, P. S., Pilgrim, D. and Riddle, D. L. (1996). Control of *C. elegans* Larval Development by Neuronal Expression of a TGF-Beta Homolog. *Science* 274, 1389-1391.

Ren, X. R., Hong, Y., Feng, Z., Yang, H. M., Mei, L. and Xiong, W. C. (2008). Tyrosine Phosphorylation of Netrin Receptors in Netrin-1 Signaling. *Neurosignals* 16, 235-245.

Rhodes, D. R., Barrette, T. R., Rubin, M. A., Ghosh, D. and Chinnaiyan, A. M. (2002). Meta-Analysis of Microarrays: Interstudy Validation of Gene Expression Profiles Reveals Pathway Dysregulation in Prostate Cancer. *Cancer Res.* 62, 4427-4433.

Ritter, A. R. and Beckstead, R. B. (2010). Sox14 is Required for Transcriptional and Developmental Responses to 20-Hydroxyecdysone at the Onset of Drosophila Metamorphosis. *Dev. Dyn.* 239, 2685-2694.

Robson, K. J., Merryweather-Clarke, A. T., Cadet, E., Viprakasit, V., Zaahl, M. G., Pointon, J. J., Weatherall, D. J. and Rochette, J. (2004). Recent Advances in Understanding Haemochromatosis: A Transition State. *J. Med. Genet.* 41, 721-730.

Rodriguez, A., Pan, P. and Parkkila, S. (2007a). Expression Studies of Neogenin and its Ligand Hemojuvelin in Mouse Tissues. *J. Histochem. Cytochem.* 55, 85-96.

Rodriguez, A., Pan, P. and Parkkila, S. (2007b). Expression Studies of Neogenin and its Ligand Hemojuvelin in Mouse Tissues. *J. Histochem. Cytochem.* 55, 85-96.

Roetto, A., Papanikolaou, G., Politou, M., Alberti, F., Girelli, D., Christakis, J., Loukopoulos, D. and Camaschella, C. (2003). Mutant Antimicrobial Peptide Hepcidin is Associated with Severe Juvenile Hemochromatosis. *Nat. Genet.* 33, 21-22.

Round, J. and Stein, E. (2007). Netrin Signaling Leading to Directed Growth Cone Steering. *Curr. Opin. Neurobiol.* 17, 15-21.

Samad, T. A., Rebbapragada, A., Bell, E., Zhang, Y., Sidis, Y., Jeong, S. J., Campagna, J. A., Perusini, S., Fabrizio, D. A., Schneyer, A. L. et al. (2005). DRAGON, a Bone Morphogenetic Protein Co-Receptor. *J. Biol. Chem.* 280, 14122-14129.

Samad, T. A., Srinivasan, A., Karchewski, L. A., Jeong, S. J., Campagna, J. A., Ji, R. R., Fabrizio, D. A., Zhang, Y., Lin, H. Y., Bell, E. et al. (2004). DRAGON: A Member of the Repulsive Guidance Molecule-Related Family of Neuronal- and Muscle-Expressed Membrane Proteins is Regulated by DRG11 and has Neuronal Adhesive Properties. *J. Neurosci.* 24, 2027-2036.

Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sun, C. Y., Baird, S. E. and Padgett, R. W. (1996). Caenorhabditis Elegans Genes Sma-2, Sma-3, and Sma-

4 Define a Conserved Family of Transforming Growth Factor Beta Pathway Components. *Proc. Natl. Acad. Sci. U. S. A.* 93, 790-794.

Savage-Dunn, C. (2005). TGF-Beta Signaling. *WormBook*, 1-12.

Savage-Dunn, C., Maduzia, L. L., Zimmerman, C. M., Roberts, A. F., Cohen, S., Tokarz, R. and Padgett, R. W. (2003). Genetic Screen for Small Body Size Mutants in *C. elegans* Reveals Many TGFbeta Pathway Components. *Genesis* 35, 239-247.

Schilham, M. W., Oosterwegel, M. A., Moerer, P., Ya, J., de Boer, P. A., van de Wetering, M., Verbeek, S., Lamers, W. H., Kruisbeek, A. M., Cumano, A. et al. (1996). Defects in Cardiac Outflow Tract Formation and Pro-B-Lymphocyte Expansion in Mice Lacking Sox-4. *Nature* 380, 711-714.

Schmierer, B. and Hill, C. S. (2007). TGFbeta-SMAD Signal Transduction: Molecular Specificity and Functional Flexibility. *Nat. Rev. Mol. Cell Biol.* 8, 970-982.

Schwartz, M. S., Benci, J. L., Selote, D. S., Sharma, A. K., Chen, A. G., Dang, H., Fares, H. and Vatamaniuk, O. K. (2010). Detoxification of Multiple Heavy Metals by a Half-Molecule ABC Transporter, HMT-1, and Coelomocytes of Caenorhabditis Elegans. *PLoS One* 5, e9564.

Serpe, M., Umulis, D., Ralston, A., Chen, J., Olson, D. J., Avanesov, A., Othmer, H., O'Connor, M. B. and Blair, S. S. (2008). The BMP-Binding Protein

Crossveinless 2 is a Short-Range, Concentration-Dependent, Biphasic Modulator of BMP Signaling in Drosophila. *Dev. Cell.* 14, 940-953.

Severyn, C. J., Shinde, U. and Rotwein, P. (2009). Molecular Biology, Genetics and Biochemistry of the Repulsive Guidance Molecule Family. *Biochem. J.* 422, 393-403.

Shah, N. and Sukumar, S. (2010). The Hox Genes and their Roles in Oncogenesis. *Nat. Rev. Cancer.* 10, 361-371.

Shemer, G. and Podbilewicz, B. (2002). LIN-39/Hox Triggers Cell Division and Represses EFF-1/fusogen-Dependent Vulval Cell Fusion. *Genes Dev.* 16, 3136-3141.

Shi, Y. and Massague, J. (2003). Mechanisms of TGF-Beta Signaling from Cell Membrane to the Nucleus. *Cell* 113, 685-700.

Siegel, P. M. and Massague, J. (2003). Cytostatic and Apoptotic Actions of TGF-Beta in Homeostasis and Cancer. *Nat. Rev. Cancer.* 3, 807-821.

Silvestri, L., Pagani, A. and Camaschella, C. (2008). Furin-Mediated Release of Soluble Hemojuvelin: A New Link between Hypoxia and Iron Homeostasis. *Blood* 111, 924-931.

Silvestri, L., Pagani, A., Fazi, C., Gerardi, G., Levi, S., Arosio, P. and Camaschella, C. (2007). Defective Targeting of Hemojuvelin to Plasma

Membrane is a Common Pathogenetic Mechanism in Juvenile Hemochromatosis. *Blood* 109, 4503-4510.

Silvestri, L., Pagani, A., Nai, A., De Domenico, I., Kaplan, J. and Camaschella, C. (2008). The Serine Protease Matriptase-2 (TMPRSS6) Inhibits Hepcidin Activation by Cleaving Membrane Hemojuvelin. *Cell. Metab.* 8, 502-511.

Smith, D. E. and Fisher, P. A. (1984). Identification, Developmental Regulation, and Response to Heat Shock of Two Antigenically Related Forms of a Major Nuclear Envelope Protein in Drosophila Embryos: Application of an Improved Method for Affinity Purification of Antibodies using Polypeptides Immobilized on Nitrocellulose Blots. *J. Cell Biol.* 99, 20-28.

Sock, E., Rettig, S. D., Enderich, J., Bosl, M. R., Tamm, E. R. and Wegner, M. (2004). Gene Targeting Reveals a Widespread Role for the High-Mobility-Group Transcription Factor Sox11 in Tissue Remodeling. *Mol. Cell. Biol.* 24, 6635-6644.

Sparkes, A. C., Mumford, K. L., Patel, U. A., Newbury, S. F. and Crane-Robinson, C. (2001). Characterization of an SRY-Like Gene, DSox14, from Drosophila. *Gene* 272, 121-129.

Srinivasan, K., Strickland, P., Valdes, A., Shin, G. C. and Hinck, L. (2003). Netrin-1/neogenin Interaction Stabilizes Multipotent Progenitor Cap Cells during Mammary Gland Morphogenesis. *Dev. Cell.* 4, 371-382.

Steiner, F. A., Talbert, P. B., Kasinathan, S., Deal, R. B. and Henikoff, S. (2012). Cell-Type-Specific Nuclei Purification from Whole Animals for Genome-Wide Expression and Chromatin Profiling. *Genome Res.* 22, 766-777.

Sulston, J. E. and Horvitz, H. R. (1977). Post-Embryonic Cell Lineages of the Nematode, Caenorhabditis Elegans. *Dev. Biol.* 56, 110-156.

Sun, L., Hui, A. M., Su, Q., Vortmeyer, A., Kotliarov, Y., Pastorino, S., Passaniti, A., Menon, J., Walling, J., Bailey, R. et al. (2006). Neuronal and Glioma-Derived Stem Cell Factor Induces Angiogenesis within the Brain. *Cancer. Cell.* 9, 287-300.

Suzuki, Y., Yandell, M. D., Roy, P. J., Krishna, S., Savage-Dunn, C., Ross, R. M., Padgett, R. W. and Wood, W. B. (1999). A BMP Homolog Acts as a Dose-Dependent Regulator of Body Size and Male Tail Patterning in Caenorhabditis Elegans. *Development* 126, 241-250. Sweetman, D. and Munsterberg, A. (2006). The Vertebrate Spalt Genes in Development and Disease. *Dev. Biol.* 293, 285-293.

Sweetman, D. and Munsterberg, A. (2006). The Vertebrate Spalt Genes in Development and Disease. *Dev. Biol.* 293, 285-293.

Takacs-Vellai, K., Vellai, T., Chen, E. B., Zhang, Y., Guerry, F., Stern, M. J. and Muller, F. (2007). Transcriptional Control of Notch Signaling by a HOX and a PBX/EXD Protein during Vulval Development in *C. elegans*. *Dev. Biol.* 302, 661-669.

Takeo, S., Akiyama, T., Firkus, C., Aigaki, T. and Nakato, H. (2005). Expression of a Secreted Form of Dally, a Drosophila Glypican, Induces Overgrowth Phenotype by Affecting Action Range of Hedgehog. *Dev. Biol.* 284, 204-218.

Talantov, D., Mazumder, A., Yu, J. X., Briggs, T., Jiang, Y., Backus, J., Atkins, D. and Wang, Y. (2005). Novel Genes Associated with Malignant Melanoma but Not Benign Melanocytic Lesions. *Clin. Cancer Res.* 11, 7234-7242.

Tanaka, S., Kamachi, Y., Tanouchi, A., Hamada, H., Jing, N. and Kondoh, H. (2004). Interplay of SOX and POU Factors in Regulation of the Nestin Gene in Neural Primordial Cells. *Mol. Cell. Biol.* 24, 8834-8846.

Tassew, N. G., Charish, J., Seidah, N. G. and Monnier, P. P. (2012). SKI-1 and Furin Generate Multiple RGMa Fragments that Regulate Axonal Growth. *Dev. Cell.* 22, 391-402.

Tassew, N. G., Chestopolava, L., Beecroft, R., Matsunaga, E., Teng, H., Chedotal, A. and Monnier, P. P. (2008). Intraretinal RGMa is Involved in Retino-Tectal Mapping. *Mol. Cell. Neurosci.* 37, 761-769.

Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., Gerald, W. L. and Massague, J. (2008). Endogenous Human microRNAs that Suppress Breast Cancer Metastasis. *Nature* 451, 147-152.

Thein, D. C., Thalhammer, J. M., Hartwig, A. C., Crenshaw, E. B.,3rd, Lefebvre, V., Wegner, M. and Sock, E. (2010). The Closely Related Transcription Factors

Sox4 and Sox11 Function as Survival Factors during Spinal Cord Development. *J. Neurochem.* 115, 131-141.

Thomas, G. (2002). Furin at the Cutting Edge: From Protein Traffic to Embryogenesis and Disease. *Nat. Rev. Mol. Cell Biol.* 3, 753-766.

Tian, C., Sen, D., Shi, H., Foehr, M. L., Plavskin, Y., Vatamaniuk, O. K. and Liu, J. (2010). The RGM Protein DRAG-1 Positively Regulates a BMP-Like Signaling Pathway in Caenorhabditis Elegans. *Development* 137, 2375-2384.

Tian, C., Shi, H., Colledge, C., Stern, M., Waterston, R. and Liu, J. (2011). The *C. elegans* SoxC Protein SEM-2 Opposes Differentiation Factors to Promote a Proliferative Blast Cell Fate in the Postembryonic Mesoderm. *Development* 138, 1033-1043.

Timmons, L. and Fire, A. (1998). Specific Interference by Ingested dsRNA. *Nature* 395, 854.

Truksa, J., Gelbart, T., Peng, H., Beutler, E., Beutler, B. and Lee, P. (2009). Suppression of the Hepcidin-Encoding Gene Hamp Permits Iron Overload in Mice Lacking both Hemojuvelin and Matriptase-2/TMPRSS6. *Br. J. Haematol.* 147, 571-581.

Truksa, J., Peng, H., Lee, P. and Beutler, E. (2006). Bone Morphogenetic Proteins 2, 4, and 9 Stimulate Murine Hepcidin 1 Expression Independently of

Hfe, Transferrin Receptor 2 (Tfr2), and IL-6. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10289-10293.

Umulis, D., O'Connor, M. B. and Blair, S. S. (2009). The Extracellular Regulation of Bone Morphogenetic Protein Signaling. *Development* 136, 3715-3728.

van de Wetering, M., Oosterwegel, M., van Norren, K. and Clevers, H. (1993). Sox-4, an Sry-Like HMG Box Protein, is a Transcriptional Activator in Lymphocytes. *EMBO J.* 12, 3847-3854.

van den Berg, D. L., Snoek, T., Mullin, N. P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I. and Poot, R. A. (2010). An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells. *Cell. Stem Cell.* 6, 369-381.

Vatamaniuk, O. K., Bucher, E. A., Sundaram, M. V. and Rea, P. A. (2005). CeHMT-1, a Putative Phytochelatin Transporter, is Required for Cadmium Tolerance in Caenorhabditis Elegans. *J. Biol. Chem*. 280, 23684-23690.

Velasco, G., Cal, S., Quesada, V., Sanchez, L. M. and Lopez-Otin, C. (2002). Matriptase-2, a Membrane-Bound Mosaic Serine Proteinase Predominantly Expressed in Human Liver and Showing Degrading Activity Against Extracellular Matrix Proteins. *J. Biol. Chem.* 277, 37637-37646.

Vielmetter, J., Kayyem, J. F., Roman, J. M. and Dreyer, W. J. (1994). Neogenin, an Avian Cell Surface Protein Expressed during Terminal Neuronal

Differentiation, is Closely Related to the Human Tumor Suppressor Molecule Deleted in Colorectal Cancer. *J. Cell Biol.* 127, 2009-2020.

Vowels, J. J. and Thomas, J. H. (1992). Genetic analysis of chemosensory control of dauer formation in Caenorhabditis elegans. *Genetics*. 130, 105-23.

Wagmaister, J. A., Miley, G. R., Morris, C. A., Gleason, J. E., Miller, L. M., Kornfeld, K. and Eisenmann, D. M. (2006). Identification of Cis-Regulatory Elements from the *C. elegans* Hox Gene Lin-39 Required for Embryonic Expression and for Regulation by the Transcription Factors LIN-1, LIN-31 and LIN-39. *Dev. Biol.* 297, 550-565.

Wakefield LM, Smith DM, Flanders KC, Sporn MB. (1988). Latent transforming growth factor-beta from human platelets. A high molecular weight complex containing precursor sequences. *J Biol Chem*. 263:7646–7654.

Wang, D., Manali, D., Wang, T., Bhat, N., Hong, N., Li, Z., Wang, L., Yan, Y., Liu, R. and Hong, Y. (2011). Identification of Pluripotency Genes in the Fish Medaka. *Int. J. Biol. Sci.* 7, 440-451.

Wang, H., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Tessier-Lavigne, M. (1999a). Netrin-3, a Mouse Homolog of Human NTN2L, is Highly Expressed in Sensory Ganglia and shows Differential Binding to Netrin Receptors. *J. Neurosci.* 19, 4938-4947.

Wang, H., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Tessier-Lavigne, M. (1999b). Netrin-3, a Mouse Homolog of Human NTN2L, is Highly Expressed in Sensory Ganglia and shows Differential Binding to Netrin Receptors. *J. Neurosci.* 19, 4938-4947.

Wang, R. H., Li, C., Xu, X., Zheng, Y., Xiao, C., Zerfas, P., Cooperman, S., Eckhaus, M., Rouault, T., Mishra, L. et al. (2005). A Role of SMAD4 in Iron Metabolism through the Positive Regulation of Hepcidin Expression. *Cell. Metab.* 2, 399-409.

Wang, X., Asplund, A. C., Porwit, A., Flygare, J., Smith, C. I., Christensson, B. and Sander, B. (2008a). The Subcellular Sox11 Distribution Pattern Identifies Subsets of Mantle Cell Lymphoma: Correlation to overall Survival. *Br. J. Haematol.* 143, 248-252.

Wang, X., Harris, R. E., Bayston, L. J. and Ashe, H. L. (2008b). Type IV Collagens Regulate BMP Signalling in Drosophila. *Nature* 455, 72-77.

Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. and Copeland, N. G. (2005). Simple and Highly Efficient BAC Recombineering using galK Selection. *Nucleic Acids Res.* 33, e36.

Wegner, M. (2010). All Purpose Sox: The Many Roles of Sox Proteins in Gene Expression. *Int. J. Biochem. Cell Biol.* 42, 381-390.

Weigle, B., Ebner, R., Temme, A., Schwind, S., Schmitz, M., Kiessling, A., Rieger, M. A., Schackert, G., Schackert, H. K. and Rieber, E. P. (2005). Highly Specific Overexpression of the Transcription Factor SOX11 in Human Malignant Gliomas. *Oncol. Rep.* 13, 139-144.

Welsh, J. B., Sapinoso, L. M., Su, A. I., Kern, S. G., Wang-Rodriguez, J., Moskaluk, C. A., Frierson, H. F.,Jr and Hampton, G. M. (2001). Analysis of Gene Expression Identifies Candidate Markers and Pharmacological Targets in Prostate Cancer. *Cancer Res.* 61, 5974-5978.

Wiater, E., Harrison, C. A., Lewis, K. A., Gray, P. C. and Vale, W. W. (2006). Identification of Distinct Inhibin and Transforming Growth Factor Beta-Binding Sites on Betaglycan: Functional Separation of Betaglycan Co-Receptor Actions. *J. Biol. Chem.* 281, 17011-17022.

Wilson, M. E., Yang, K. Y., Kalousova, A., Lau, J., Kosaka, Y., Lynn, F. C., Wang, J., Mrejen, C., Episkopou, V., Clevers, H. C. et al. (2005). The HMG Box Transcription Factor Sox4 Contributes to the Development of the Endocrine Pancreas. *Diabetes* 54, 3402-3409.

Wilson, N. H. and Key, B. (2006). Neogenin Interacts with RGMa and Netrin-1 to Guide Axons within the Embryonic Vertebrate Forebrain. *Dev. Biol.* 296, 485-498.

Wilson, N. H. and Key, B. (2007). Neogenin: One Receptor, Many Functions. *Int. J. Biochem. Cell Biol.* 39, 874-878.

Xia, Y., Babitt, J. L., Sidis, Y., Chung, R. T. and Lin, H. Y. (2008). Hemojuvelin Regulates Hepcidin Expression Via a Selective Subset of BMP Ligands and Receptors Independently of Neogenin. *Blood* 111, 5195-5204.

Xia, Y., Cortez-Retamozo, V., Niederkofler, V., Salie, R., Chen, S., Samad, T. A., Hong, C. C., Arber, S., Vyas, J. M., Weissleder, R. et al. (2011). Dragon (Repulsive Guidance Molecule b) Inhibits IL-6 Expression in Macrophages. *J. Immunol.* 186, 1369-1376.

Xia, Y., Yu, P. B., Sidis, Y., Beppu, H., Bloch, K. D., Schneyer, A. L. and Lin, H. Y. (2007). Repulsive Guidance Molecule RGMa Alters Utilization of Bone Morphogenetic Protein (BMP) Type II Receptors by BMP2 and BMP4. *J. Biol. Chem.* 282, 18129-18140.

Xie, Y., Ding, Y. Q., Hong, Y., Feng, Z., Navarre, S., Xi, C. X., Zhu, X. J., Wang, C. L., Ackerman, S. L., Kozlowski, D. et al. (2005). Phosphatidylinositol Transfer Protein-Alpha in Netrin-1-Induced PLC Signalling and Neurite Outgrowth. *Nat. Cell Biol.* 7, 1124-1132.

Xie, Y., Hong, Y., Ma, X. Y., Ren, X. R., Ackerman, S., Mei, L. and Xiong, W. C. (2006). DCC-Dependent Phospholipase C Signaling in Netrin-1-Induced Neurite Elongation. *J. Biol. Chem.* 281, 2605-2611.

Yamashita, T., Mueller, B. K. and Hata, K. (2007). Neogenin and Repulsive Guidance Molecule Signaling in the Central Nervous System. *Curr. Opin. Neurobiol.* 17, 29-34.

Yang, F., West, A. P.,Jr, Allendorph, G. P., Choe, S. and Bjorkman, P. J. (2008). Neogenin Interacts with Hemojuvelin through its Two Membrane-Proximal Fibronectin Type III Domains. *Biochemistry* 47, 4237-4245.

Yang, F., West, A. P.,Jr and Bjorkman, P. J. (2011). Crystal Structure of a Hemojuvelin-Binding Fragment of Neogenin at 1.8A. *J. Struct. Biol.* 174, 239-244.

Yang, J., Chai, L., Fowles, T. C., Alipio, Z., Xu, D., Fink, L. M., Ward, D. C. and Ma, Y. (2008). Genome-Wide Analysis Reveals Sall4 to be a Major Regulator of Pluripotency in Murine-Embryonic Stem Cells. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19756-19761.

Yang, J., Gao, C., Chai, L. and Ma, Y. (2010). A Novel SALL4/OCT4 Transcriptional Feedback Network for Pluripotency of Embryonic Stem Cells. *PLoS One* 5, e10766.

Yang, J., Liao, W. and Ma, Y. (2012). Role of SALL4 in Hematopoiesis. *Curr. Opin. Hematol.*

Yuri, S., Fujimura, S., Nimura, K., Takeda, N., Toyooka, Y., Fujimura, Y., Aburatani, H., Ura, K., Koseki, H., Niwa, H. et al. (2009). Sall4 is Essential for Stabilization, but Not for Pluripotency, of Embryonic Stem Cells by Repressing Aberrant Trophectoderm Gene Expression. *Stem Cells* 27, 796-805.

Zhang, A. S. (2010). Control of Systemic Iron Homeostasis by the Hemojuvelin-Hepcidin Axis. *Adv. Nutr.* 1, 38-45.

Zhang, A. S., Anderson, S. A., Meyers, K. R., Hernandez, C., Eisenstein, R. S. and Enns, C. A. (2007). Evidence that Inhibition of Hemojuvelin Shedding in Response to Iron is Mediated through Neogenin. *J. Biol. Chem.* 282, 12547-12556.

Zhang, A. S., Gao, J., Koeberl, D. D. and Enns, C. A. (2010). The Role of Hepatocyte Hemojuvelin in the Regulation of Bone Morphogenic Protein-6 and Hepcidin Expression *in vivo*. *J. Biol. Chem.* 285, 16416-16423.

Zhang, A. S., West, A. P.,Jr, Wyman, A. E., Bjorkman, P. J. and Enns, C. A. (2005). Interaction of Hemojuvelin with Neogenin Results in Iron Accumulation in Human Embryonic Kidney 293 Cells. *J. Biol. Chem.* 280, 33885-33894.

Zhang, A. S., Yang, F., Meyer, K., Hernandez, C., Chapman-Arvedson, T., Bjorkman, P. J. and Enns, C. A. (2008). Neogenin-Mediated Hemojuvelin Shedding Occurs After Hemojuvelin Traffics to the Plasma Membrane. *J. Biol. Chem.* 283, 17494-17502.

Zhang, A. S., Yang, F., Wang, J., Tsukamoto, H. and Enns, C. A. (2009). Hemojuvelin-Neogenin Interaction is Required for Bone Morphogenic Protein-4-Induced Hepcidin Expression. *J. Biol. Chem.* 284, 22580-22589.

Zhang, J. L., Qiu, L. Y., Kotzsch, A., Weidauer, S., Patterson, L., Hammerschmidt, M., Sebald, W. and Mueller, T. D. (2008). Crystal Structure Analysis Reveals how the Chordin Family Member Crossveinless 2 Blocks BMP-2 Receptor Binding. *Dev. Cell.* 14, 739-750.

Zhao, Z. W., Lian, W. J., Chen, G. Q., Zhou, H. Y., Wang, G. M., Cao, X., Yang, H. J. and Hou, Y. P. (2012). Decreased Expression of Repulsive Guidance Molecule Member A by DNA Methylation in Colorectal Cancer is Related to Tumor Progression. *Oncol. Rep.* 27, 1653-1659.

Zhou, Z., Xie, J., Lee, D., Liu, Y., Jung, J., Zhou, L., Xiong, S., Mei, L. and Xiong, W. C. (2010). Neogenin Regulation of BMP-Induced Canonical Smad Signaling and Endochondral Bone Formation. *Dev. Cell.* 19, 90-102.